



# A NONLINEAR PERTURBATION THEORY FOR ESTIMATION AND CONTROL OF TIME-DISCRETE STOCHASTIC SYSTEMS

H. W. Sorenson

January 1968  
Report No. 68-2

FACILITY FORM 602

N68-16918

(ACCESSION NUMBER)

244  
(PAGES)

CR-92981  
(NASA CR OR TMX OR AD NUMBER)

(THRU)

(CODE)

10  
(CATEGORY)

**A NONLINEAR PERTURBATION THEORY FOR ESTIMATION AND CONTROL  
OF TIME-DISCRETE STOCHASTIC SYSTEMS**

H. W. SORENSON

DEPARTMENT OF ENGINEERING  
UNIVERSITY OF CALIFORNIA  
LOS ANGELES

## FOREWORD

The research described in this report, "A Nonlinear Perturbation for Estimation and Control of Time-Discrete Stochastic Systems," Number 68-2, by H. W. Sorenson, was carried out under the direction of C. T. Leondes in the Department of Engineering, University of California, Los Angeles.

This project is supported in part by the National Aeronautics and Space Administration Grant NsG 237-62 to the Institute of Geophysics and Interplanetary Physics of the University.

This report is based on a Doctor of Philosophy dissertation submitted by the author.

## TABLE OF CONTENTS

LIST OF FIGURES	vi
LIST OF TABLES	vii
NOMENCLATURE	viii
ABSTRACT	xi
CHAPTER ONE - GENERAL DISCUSSION AND PROBLEM STATEMENT	1
1.1 The General Problem	2
1.2 Previous Results	9
1.3 Preview of Coming Attractions	13
CHAPTER TWO - THE BAYESIAN APPROACH	19
2.1 Optimal Estimation and Control for Stochastic Time-Discrete Systems	20
2.1.1 <u>The Minimum Mean-Square Estimation Problem</u>	20
2.1.2 <u>The Control Problem</u>	22
2.2 The A Posteriori Conditional Density Function	27
2.2.1 <u>Recursion Relation for <math>p(\underline{x}_k/\underline{z}^k)</math></u>	27
2.2.2 <u>The A Posteriori Density Function for Prediction and Smoothing</u>	30
2.3 Characteristic Function Equivalents	33
CHAPTER THREE - THE LINEAR, TIME-DISCRETE STOCHASTIC CONTROL PROBLEM	39
3.1 Minimum Mean-Square Estimates	40
3.2 The Linear Feedback Control Law	50

CHAPTER FOUR - A GENERALIZATION OF THE KALMAN FILTER	57
4.1 An A Posteriori Gaussian Density for Filtering of Non-linear Systems	58
4.2 On the Approximation of Nonlinear Systems	66
4.2.1 <u>Relation to the Kalman Filter</u>	67
4.2.2 <u>Choice of the Nominal</u>	69
4.2.3 <u>Conclusions</u>	71
4.3 On the Control of a Linear Plant using Nonlinear Measurement Data	72
CHAPTER FIVE - FILTERING FOR NAVIGATION OF A SPACECRAFT	77
5.1 The Space Navigation Problem	77
5.1.1 <u>The Mathematical Model</u>	78
5.1.2 <u>The Estimation Policies</u>	81
5.2 Numerical Results	87
5.2.1 <u>When Linear Theory is Valid</u>	89
5.2.2 <u>Orbit Rectification to the Rescue</u>	92
5.3 Conclusions of the Computational Study	101
CHAPTER SIX - APPROXIMATION OF THE A POSTERIORI DENSITY FUNCTION FOR NONLINEAR SYSTEMS	103
6.1 The Approximation Procedure	104
6.2 The Conditional Density for a Second Order System	109
6.3 On Extending the Use of Linear Filters	118
CHAPTER SEVEN - NUMERICAL COMPARISON OF LINEAR AND NONLINEAR FILTERING TECHNIQUES	129
7.1 The General Problem	130

7.2	Filters Based Upon a Gaussian Density	132
7.3	Estimators Based Upon Nongaussian Densities	142
7.3.1	<u>Nonlinear, Nongaussian Filters</u>	143
7.3.2	<u>A Poor Density Approximation</u>	148
7.4	Summary of Results	151
CHAPTER EIGHT - SUMMARY AND CONCLUDING REMARKS		155
8.1	Summary	155
8.2	General Conclusions and Suggestions for Future Research	161
REFERENCES		167
APPENDIX A - GAUSSIAN A POSTERIORI DENSITY FUNCTION		173
APPENDIX B - MATHEMATICAL MODEL FOR THE SPACE NAVIGATION PROBLEM		185
APPENDIX C - GRAM-CHARLIER AND EDGEWORTH EXPANSIONS		195
APPENDIX D - DERIVATION OF NONGAUSSIAN A POSTERIORI DENSITY FUNCTION		205
APPENDIX E - MOMENTS OF A DISTRIBUTION		227

## LIST OF FIGURES

5.1	Error in Linear Approximation - Case 1	91
5.2	Error in Linear Approximation - Case 2	96
7.1(a)	Comparison of Gaussian Estimators - Noise-free Plant	135
7.1(b)	Comparison of Gaussian Estimators - Noise-free Plant	137
7.1(c)	Comparison of Gaussian Estimators - Noise-free Plant	138
7.2	Comparison of Gaussian Estimators - Noisy Plant	140
7.3	Comparison of Gaussian Estimators - Large Initial Perturbation	141
7.4(a)	Comparison of Gaussian and Nongaussian Estimators - Noise-free Plant	145
7.4(b)	Comparison of Gaussian and Nongaussian Estimators - Noise-free Plant	146
7.4(c)	Comparison of Gaussian and Nongaussian Estimators - Noise-free Plant	147
7.5(a)	Filter Response for Large Initial Perturbations	149
7.5(b)	Filter Response for Large Initial Perturbations	150
B-1	Geometry of Angular Measurements	189
C-1	Series Approximations to Uniform Probability Density	202
C-2	Approximation of $\chi^2$ -distribution for $n = 1$	203
C-3	Approximation of $\chi^2$ -distribution for $n = 4$	204

## LIST OF TABLES

5.1(a)	Filter Response when the Perturbations are Linear	93
5.1(b)	Filter Response when the Perturbations are Linear	94
5.2(a)	Filter Response when a Prespecified Nominal is Inadequate	98
5.2(b)	Filter Response when a Prespecified Nominal is Inadequate	99
6.1	The A Posteriori Density Function	122-127



## NOMENCLATURE

$\underline{a}$	expected value of the initial state
$B(\underline{x}_k)$	quantity describing the nonlinear measurement effects when expressed in terms of the state $\underline{x}_k$
$B_1(\eta_k)$	quantity describing the nonlinear measurement effects when expressed in terms of the centered variable $\eta_k = \underline{x}_k - \hat{\xi}_k$ . The $\hat{\xi}_k$ represents the conditional mean provided by a linear approximation.
$\underline{f}_k$	function that relates past and current state: $\underline{x}_k = \underline{f}_k(\underline{x}_{k-1}, \underline{u}_{k-1}, \underline{w}_{k-1})$
$\underline{h}_k$	function that relates measurement and state: $\underline{z}_k = \underline{h}_k(\underline{x}_k, \underline{v}_k)$
$H_k$	linear observation matrix containing the first partial derivatives of $\underline{h}_k$ with respect to $\underline{x}_k$
$H_i(\ )$	the $i^{\text{th}}$ Hermite polynomial
$\mathcal{H}_{N-k}$	criterion for establishing the optimal control at $t_{N-k}$ . The optimal control is chosen to minimize the conditional expectation of $\mathcal{H}_{N-k}$
$J_k^i$	the second order observation matrix containing the second partial derivatives of $\underline{h}_k$ with respect to the state $\underline{x}_k$
$M_o$	covariance matrix for the initial state
$P_{k/k+\gamma}$	covariance matrix of the a posteriori density function of the state $\underline{x}_k$ conditioned on the measurements $\underline{z}^{k+\gamma}$

$P_k$	shorthand notation for $P_{k/k}$
$P'_k$	shorthand notation for $P_{k+1/k}$
$Q_k$	covariance matrix for the plant noise $\underline{w}_k$
$R_k$	covariance matrix for the measurement noise $\underline{v}_k$
$t_k$	time at the $k^{\text{th}}$ sampling interval ( $k = 0, 1, \dots, N$ )
$\underline{u}_k$	p-dimensional control vector for the interval $[t_k, t_{k+1})$
$\underline{v}_k$	white noise sequence corrupting the measurement data
$V_N$	performance index for N control intervals
$\underline{w}_k$	white noise sequence acting upon plant
$W_i(\underline{x}_k, \underline{u}_{k-1})$	weighting function in performance index: $V_N = \sum_{i=1}^N W_i(\underline{x}_i, \underline{u}_{i-1})$
$W_i^X, W_{i-1}^U$	Arbitrary weighting matrices in quadratic performance index: $V_N = \sum_{i=1}^N (\underline{x}_i^T W_i^X \underline{x}_i + \underline{u}_{i-1}^T W_{i-1}^U \underline{u}_{i-1})$
$\underline{x}_k$	n-dimensional state vector
$\hat{\underline{x}}_{k/k+\gamma}$	mean of the a posteriori density function of the state $\underline{x}_k$ conditioned on the measurements $\underline{z}^{k+\gamma}$
$\hat{\underline{x}}_k$	shorthand notation for $\hat{\underline{x}}_{k/k}$
$\hat{\underline{x}}'_k$	shorthand notation for $\hat{\underline{x}}_{k+1/k}$
$\tilde{\underline{x}}_{k/k+\gamma}$	error in the estimate $\hat{\underline{x}}_{k/k+\gamma}$ : $\tilde{\underline{x}}_{k/k+\gamma} \stackrel{\text{Df}}{=} \hat{\underline{x}}_{k/k+\gamma} - \underline{x}_k$
$\underline{z}_k$	m-dimensional vector of measurement data
$\mu_k$	third central moment of the a posteriori density function $p(\underline{x}_k / \underline{z}^k)$

$\nu_k$	fourth central moment of the a posteriori density function
$p(x_k/z^k)$	
$I$	identity matrix
$\delta_{kj}$	Kronecker delta
$\delta(\ )$	Dirac delta function
$\exp(\ )$	exponential function
$(\ )_k$	the quantity $(\ )$ at $t_k$
$(\_)$	the quantity $(\_)$ is a vector
$(\_)^k$	the collection $(\_)_1, (\_)_2, \dots, (\_)_k$
$(\ )^*$	the nominal value of $(\ )$
$(\ )^T$	the matrix transpose of $(\ )$
$(\ )^{-1}$	the matrix inverse of $(\ )$
$p(\ )$	the probability density function of $(\ )$
$p(a/b)$	the conditional probability density function of a given b
$\varphi(\ )$	the characteristic function of $(\ )$
$E(\ )$	the expected value of $(\ )$
$E(a/b)$	the expected value of a given b
$(\frac{\partial \underline{f}}{\partial \underline{x}})$	the matrix of partial derivatives of the components $f^i$ of $\underline{f}$ taken with respect to the components $x^i$ of $\underline{x}$
$(\frac{\partial^2 f^i}{\partial x^j \partial x^k})$	the matrix of second partial derivatives of the $i^{\text{th}}$ component of $\underline{f}$ taken with respect to the components $x^j$ of $\underline{x}$
$a \stackrel{\text{Df}}{=} b$	the quantity a is defined to be equivalent to b

## ABSTRACT

The problem of determining optimal estimation and control policies from noisy measurement data for time-discrete, stochastic, dynamical systems is considered in this dissertation. The method that is proposed here for the solution of these problems represents a generalization of the common approach that is based on the application of linear theory. In applying linear theory, it is assumed that the state and measurement perturbations of the actual system relative to an arbitrary system can be described by linear equations. Then, it is possible to apply well-known linear techniques to estimate the state perturbations and to determine the desired control corrections. In this investigation, terms of higher order than first are retained in describing the perturbations. The determination of estimation and control policies for the resulting nonlinear systems is then accomplished within the framework of the so-called Bayesian approach.

The general solution of the estimation and control problems can be established if the a posteriori density function  $p(\underline{x}_k / \underline{z}^k)$  of the state conditioned upon all past and current measurement data is known. It is not possible to express this density in a closed-form in most cases, so a principal concern in this study is with the approximation, rather than the precise determination, of the  $p(\underline{x}_k / \underline{z}^k)$ . A general procedure for approximating the densities is proposed and then applied to a specific nonlinear system. For this system, the plant and measurement noise is assumed to be additive and gaussian. Then, the a posteriori density is approximated by a truncated Edgeworth expansion that includes the fourth central moments. Using this form for the approximation,

recurrence relations for the moments of the distribution are developed.

These equations can be simplified in a straightforward manner to yield several other approximations. This includes a gaussian approximation that is more general than the results obtained by first assuming a linear model.

The estimation problem was considered in some detail. Techniques are suggested that allow the range of applicability of linear theory to be considerably extended. This extension is illustrated by numerical examples in which the estimates obtained from the standard Kalman filter, modified Kalman filters, and the nonlinear filters are compared. The proposed modifications are seen to yield significant improvements in many cases. These results suggest that for many problems it might be fruitful to explore these and other modified linear techniques before attempting to apply a nonlinear theory. However, problems do exist that require the use of nonlinear methods. The approach suggested here leads to results that are reasonable for use with digital computers and appears to warrant further investigation.

## CHAPTER ONE

### GENERAL DISCUSSION AND PROBLEM STATEMENT

The Austrian physicist Ludwig Boltzmann is reputed to have once remarked that "there is nothing more practical than a good theory". Believing this aphorism to be a worthy engineering watchword, it is the intent in this study to investigate the problem of establishing estimation and control policies for stochastic dynamical systems by considering a general theory, namely, the so-called Bayesian approach. As with many such pithy statements, one or more words can be subject to diverse interpretation. In Boltzmann's phrase, the key word would appear to be "good", and we suggest that for many engineers, it might be defined in the following, almost circular, manner. A theory is good if it leads to the understanding and solution, either analytically or numerically, of practical problems. Thus, after formulating the general problem and theory in Chapters 1 and 2, considerable emphasis is placed upon the application of the theory and the development of computational algorithms.

In Section 1.1, the mathematical model and the general problem are stated and many of the terms and notations that appear throughout the text are introduced and discussed. Results that have appeared in the literature relative to the general topic considered in this study are reviewed in Section 1.2. This discussion can by no means be considered to exhaust the subject. Additional references appear throughout the text. In the final section of this chapter, the theory that is proposed here for the solution of the estimation and control problems is presented. Also, an outline of the contents of Chapters 2 through 8 is provided.

## 1.1 THE GENERAL PROBLEM

As has been stated, the problem of determining estimation and control policies for stochastic dynamical systems is to be considered. Before stating these problems, several terms need to be defined.

First, it is important to recognize the precise meaning of "stochastic" as used here. Certainly, it implies the probabilistic nature of the investigation but, moreover, we use it to imply that the a priori distributions of all random quantities are completely known. In this sense, we follow Bellman [1,2] who has suggested that a system be called adaptive when parameters of the distributions are unknown and must be "learned". This is in contrast to the case in which a parameter of the dynamical system is unknown but has an a priori distribution that is completely defined. This example would still be a stochastic problem, although the parameter must be estimated (or learned).

Only N-stage, time-discrete systems are considered in the ensuing discussion. In general, the state [3]  $\underline{x}_k$  of the dynamical system is assumed to evolve according to the nonlinear difference equation

$$\underline{x}_k = \underline{f}_k(\underline{x}_{k-1}, \underline{u}_{k-1}, \underline{w}_{k-1}) \quad k = 1, \dots, N \quad (I)$$

where the state  $\underline{x}_k$  is n-dimensional. The p-dimensional vector  $\underline{u}_{k-1}$  describes the control parameters that are to be selected according to a prescribed control law. At each time, the system is disturbed by the random noise  $\underline{w}_{k-1}$ . Throughout the discussion, the sequence\*  $\underline{w}^{k-1}$  is assumed to have a known

---

\* The notation  $\underline{a}^k$  is used to designate the collection  $(\underline{a}_0, \underline{a}_1, \dots, \underline{a}_k)$ .

probability density  $p(\underline{w}^{k-1})$  and to be independent from one sampling time to the next. That is,

$$p(\underline{w}_0, \underline{w}_1, \dots, \underline{w}_k) \stackrel{\text{Df}}{=} p(\underline{w}^k) = p(\underline{w}_0)p(\underline{w}_1) \dots p(\underline{w}_k)$$

Sequences having this characteristic shall be referred to as white noise sequences (not to be confused with white noise processes which have a considerably different character).

The notation that is used follows Fel'dbaum [4-9] and has the disadvantage that the argument of a function serves a dual purpose. It is used to name the function (as is done above) and is also treated as a variable name (e.g., it is treated as the variable of integration). The meaning should be clear from the context.

The initial condition for the state  $\underline{x}_0$  is also a random variable with a known probability density  $p(\underline{x}_0)$ . Note that the probability density is assumed to exist in this and all other examples. This does not represent a significant restriction and could be replaced in each instance by the Radon-Nikodym derivative. The  $\underline{x}_0$  is assumed to be independent of the noise sequences in the plant and measurements.

The function  $\underline{f}_k$  in (I) is considered to be known. This relation is frequently referred to as the plant equation and the fixed system that defines  $\underline{f}_k$  as the plant.

The behavior of the plant is generally observed imperfectly through the measurement of quantities  $\underline{z}_k$  that are functionally related to the state variables and which contain random errors. These data are assumed to be described by the known relation



$$\underline{z}_k = \underline{h}_k(\underline{x}_k, \underline{v}_k) \quad k = 0, 1, \dots, N-1 \quad (\text{II})$$

where  $\underline{z}_k$  is a  $m$ -dimensional vector. The noise  $\underline{v}_k$  is supposed to be a member of a white noise sequence with known density  $p(\underline{v}_k)$ .

Equations (I) and (II) constitute the basic mathematical model for the study. Note that equations that are deemed to be of particular importance shall be denoted with the Roman numeral as has been done for (I) and (II). Arabic symbols shall be used for equations having a more secondary nature. The subscripts E and C will be used for equations that are significant for either the estimation or the control problem, but not both.

It is now possible to give a more explicit definition of the estimation and control problems.

**ESTIMATION:** The estimation problem is essentially concerned with the determination of the state  $\underline{x}_k$  from the measurement data  $\underline{z}^{k+\gamma}$ . The problem separates naturally into three subproblems.

- 1) Filtering: estimate  $\underline{x}_k$  from all past and current measurement data  $\underline{z}^k$  (i. e. ,  $\gamma = 0$ )
- 2) Prediction: predict  $\underline{x}_k$  from past data (i. e. ,  $\gamma < 0$ )
- 3) Smoothing: estimate  $\underline{x}_k$  using future data as well as past and current data (i. e. ,  $\gamma > 0$ ).

All three cases shall be dealt with in the succeeding pages, but the greatest emphasis is placed upon the filtering problem. In particular, we shall consider the recursive filtering problem in which the estimate  $\hat{\underline{x}}_k$  shall be based upon  $\hat{\underline{x}}_{k-1}$  and  $\underline{z}_k$ .

Because of the presence of noise in the plant and measurement equations, it is, in general, not possible to determine  $\underline{x}_k$  precisely from the data  $\underline{z}^{k+\gamma}$ . Instead, the estimate  $\hat{\underline{x}}_{k/k+\gamma}$  must be chosen to approximate  $\underline{x}_k$  in some well-defined sense. Suppose the error in the estimate is denoted as  $\tilde{\underline{x}}_{k/k+\gamma}$  and is defined as

$$\tilde{\underline{x}}_{k/k+\gamma} \stackrel{\text{Df}}{=} \hat{\underline{x}}_{k/k+\gamma} - \underline{x}_k$$

The error criteria that is selected generally has the form of  $E\{\varphi(\tilde{\underline{x}}_{k/k+\gamma})\}$  where  $\varphi$  is positive and spherically symmetric. That is, it is true that

$$0 \leq \varphi(\tilde{\underline{x}}_{k/k+\gamma}) = \varphi(-\tilde{\underline{x}}_{k/k+\gamma})$$

and such that if

$$0 \leq \left| \tilde{\underline{x}}_{k/k+\gamma}^{(1)} \right| \leq \left| \tilde{\underline{x}}_{k/k+\gamma}^{(2)} \right|$$

then

$$\varphi(\tilde{\underline{x}}_{k/k+\gamma}^{(1)}) \leq \varphi(\tilde{\underline{x}}_{k/k+\gamma}^{(2)})$$

Examples of error criteria that satisfy these conditions are:

- 1) Minimum mean-square error.

For this criteria, the estimate is chosen to cause

$$E[\tilde{\underline{x}}_{k/k+\gamma}^T \tilde{\underline{x}}_{k/k+\gamma}] = \text{minimum.}$$

- 2) Minimum absolute deviation.

In this case the estimate is chosen so that

$$E[|\tilde{\underline{x}}_{k/k+\gamma}|] = \text{minimum.}$$

It is well-known and will be demonstrated in Section 2.1 that the mean square error is minimized by choosing  $\hat{\underline{x}}_{k/k+\gamma}$  to be the mean of the conditional density  $p(\underline{x}_k/\underline{z}^{k+\gamma})$ . It is also known [10] that in the scalar case the minimum absolute deviation is obtained by choosing the estimate to be the median of  $p(\underline{x}_k/\underline{z}^{k+\gamma})$ .

For the scalar case, Sherman [11] has pointed out the following lemma.

LEMMA: For the  $\varphi$  defined above and if  $P$  is a probability distribution on the reals which is symmetric and unimodal with mode at the origin so that  $P(X) = 1 - P(-X)$  at each continuity point of  $P$  and  $P$  is convex for  $X < 0$ , then

$$\int \varphi(X) dP(X) \leq \int \varphi(X - a) dP(X)$$

for each real  $a$ , when the integrals exist; if either integral diverges, the one on the right does.

This implies that for conditional distributions satisfying the conditions of the lemma, the estimate for error criteria  $E[\varphi(\tilde{\underline{x}}_{k/k+\gamma})]$  is the same as for the minimum mean-square error criteria. Thus, estimates based on the latter criteria can encompass a much larger class than is popularly believed. For the remainder of this discussion, only mean-square estimates shall be considered. Thus, the estimates will be selected to minimize

$$\varphi(\tilde{\underline{x}}_{k/k+\gamma}) = E[\tilde{\underline{x}}_{k/k+\gamma}^T \tilde{\underline{x}}_{k/k+\gamma}] \quad (\text{III}_E)$$

An additional criteria for selecting the estimate would be to select  $\hat{\underline{x}}_{k/k+\gamma}$  as the maximum value of the conditional density function  $p(\underline{x}_k/\underline{z}^{k+\gamma})$ .

This is sometimes referred to as the "most probable" estimate and is the mode of the distribution. Cox [12] has considered this estimate in considerable detail.

It has the disadvantage that there is no natural measure of error to attribute to the estimate.

CONTROL: The plant (I) is caused to behave in a particular manner through the selection of the control vectors  $\underline{u}^{N-1}$ . The rule according to which the  $\underline{u}_k$  are selected at each sampling time ( $k = 0, 1, \dots, N-1$ ) is called the control law for the system. As was true in the choice of estimates discussed above, the means of establishing the control law is somewhat arbitrary. In the following, we shall assume that the control is chosen to minimize the expected value of the performance index

$$V_N = \sum_{i=1}^N W_i(\underline{x}_i, \underline{u}_{i-1}) \quad (\text{III}_C)$$

The  $W_i$  are specified functions of the state and control variables and shall be required to be nonnegative and spherically symmetric. A familiar example and one that will be used later is the quadratic index

$$V_N = \sum_{i=1}^N (\underline{x}_i^T W_i^X \underline{x}_i + \underline{u}_{i-1}^T W_{i-1}^U \underline{u}_{i-1})$$

where the  $W_i^X$  and  $W_{i-1}^U$  are arbitrary, non-negative definite weighting matrices.

The behavior of the system is observed through the measurement data  $\underline{z}^k$  so the control law is taken as a function of these data. That is, at each sampling time  $t_k$ ,  $t_0 \leq t_k \leq t_{N-1}$ , the control is computed according to

$$\underline{u}_k = \underline{U}_k [\underline{z}^{k-\gamma}]$$

The  $\gamma$  has been included to indicate that the control might be based on past data only. Physical realizability considerations require that  $\gamma \geq 0$  since the control could not be expected to depend upon future measurements.

It would appear that a more general control law could be obtained if  $\underline{u}_k$  were allowed to be a random (rather than deterministic) function of the measurement data. Fel'dbaum considered this possibility and found [5] that the generalization did not provide any benefit in the cases that he considered. Sworder [13, 14] has shown that it is sufficient to consider deterministic control laws for Bayesian control policies.

The form of the optimal control law for a given system (I) and performance index ( $III_C$ ) depends upon the nature of the observational information that is assumed to be available to the controller. The two conditions that are of greatest interest occur when  $\gamma = 0$  and when  $\gamma = k$ . The former results in a feedback (or closed-loop) control law, whereas the latter leads to an open-loop control law. In deterministic problems, there is no difference between the two types of control.

In open-loop control, the entire control policy is established by the initial conditions, whether this is represented by  $\underline{x}_0$  or measurements made prior to the initiation of control. This can be modified to a policy that has been referred to as an open-loop feedback control law. In this case, the control policy is computed anew at each  $t_k$  by treating  $t_k$  as the initial time and by ignoring the fact that new data will be available at later times. Open-loop, feedback control might be expected to produce a policy that is superior to open-loop control but inferior to feedback control. Dreyfus [15] demonstrated that this intuitive idea is valid for a simple stochastic control problem. Katz [16] shows that the feedback policy provides a lower bound for the value of the performance index when the systems are time-continuous.

A fourth alternative has been suggested by Simon [17] and has been called a certainty equivalence control policy. In this case the random variables are replaced by their unconditional mean values and the problem is treated as deterministic. This policy has been shown to provide a solution to the stochastic control problem for linear systems containing white noise sequences and with a quadratic performance index. This situation is discussed in Chapter 3.

The problem of determining feedback control policies is considered in Chapters 2 and 3.

## 1.2 PREVIOUS RESULTS

Research into the stochastic control problem has quite naturally taken two avenues of approach. In the preceding section, the problem has been posed in terms of a time-discrete system involving difference equations and a finite number of observation and control times. It could reasonably have been stated instead in terms of a time-continuous system with a differential equation model and continuous measurement and control processes. Since dynamical systems are usually described by differential equations, it could be concluded that this would be the more natural model. A considerable amount of research effort has been expended in this area. For a summary, see References 18 or 19. More recent results than those described in the aforementioned references have been published by Bucy [20], Bass [21], Mortensen [22], and Fisher [23]. The first two have dealt with the estimation problem, whereas Mortensen has presented a very general and mathematically sophisticated solution of the control problem. Fisher considered the estimation problem

from the point of view of approximating the a posteriori density function of a time-continuous system.

There are advantages and disadvantages to both formulations. The principal disadvantage of the time-discrete model arises from the fact that, as has already been mentioned, a dynamical system is generally described by a system of differential equations. In order to obtain the time-discrete model, it is necessary to reduce the system to the form described by (I) and this requirement engenders a problem of considerable significance. On the other hand, it is believed that the formulation presented in Section 1.1 is more realistic for several reasons.

- (1) Measurement data are usually available only at discrete times.
- (2) In many complex systems, the control is determined with the aid of digital computer so the control is changed at discrete times.
- (3) In the time-continuous case, white noise processes are generally assumed to act on the plant and measurement process and such noise is physically unrealizable.
- (4) Last, and not least, the general solution of the time-continuous estimation and control problems yields systems of complicated partial differential-integral equations that must be solved. The difficulties inherent in obtaining numerical solutions to practical problems using this formulation appear to be excessive.

When (I) and (II) are linear, the noise is Gaussian, and a minimum mean-square error criteria and a quadratic performance index are utilized, the solutions to the estimation and control problem are well-established. There have been many workers in this area, but many of the better known results have been attributed to R. E. Kalman [24,25,26]. It was suggested by Kalman and Koepcke [27] that for linear systems the estimation and control problems could be considered separately. That is, the estimates can be computed as though the control is a known function of time and the control law found for the deterministic problem can be used for the stochastic control law. The control is computed according to

$$\underline{u}_k = \Lambda_k \hat{\underline{x}}_k$$

where  $\Lambda_k$  describes the deterministic control law, and  $\hat{\underline{x}}_k$  has replaced  $\underline{x}_k$ . This result has been stated as a "Separation Theorem" and was first proven independently by Gunckel [28] and by Joseph [29].

For nonlinear systems, an approach that is commonly used in practice involves the use of linear perturbation techniques [30]. First, a nominal or reference solution of (I) is assumed to exist that provides a "good" approximation to the actual behavior of the system. The approximation is "good" if the difference  $\delta \underline{x}$  between the nominal and actual states can be accurately described by a system of linear difference equations

$$\delta \underline{x}_k = \Phi_{k,k-1} \delta \underline{x}_{k-1} + \Gamma_{k,k-1} \underline{u}_{k-1} + \Delta_{k,k-1} \underline{w}_{k-1}$$

and the difference in the measurements  $\delta \underline{z}_k$  is given by



$$\delta \underline{z}_k = H_k \delta \underline{x}_k + \underline{v}_k$$

This approach has yielded many satisfactory results, but several weaknesses have become apparent. For example,

- (1) There is no easily obtained criteria for judging the validity of the linear approximations.
- (2) The filter does not behave satisfactorily when the measurement noise is small. Pines and Denham [31] have attributed this to the absence of second order terms in the expansion of the measurement equations.
- (3) This procedure lacks generality. It provides little insight into the techniques for considering more general systems.

It has been suggested by several people that it would be more appropriate to formulate the problem in terms of the a posteriori density function  $p(\underline{x}_k/\underline{z}^k)$ . In a series of four papers, Fel'dbaum [4 - 7] dealt with the control problem and derived several basic results. Ho and Lee [32] considered the estimation problem. Aoki [33] has conducted an extensive investigation of both problems. These results are contained in his forthcoming book. In an excellent doctoral dissertation, Sworder [13,14] has considered the control problem using a game-theoretic formulation. Stratonovich [34] dealt with the a posteriori density for time-discrete and time-continuous systems.

Knowledge of  $p(\underline{x}_k/\underline{z}^k)$  theoretically provides the solution to both the estimation and control problems. The estimates  $\hat{\underline{x}}_k$  and control  $\underline{u}_k$  are required at each sampling instant so it is necessary to know  $p(\underline{x}_k/\underline{z}^k)$  for

every  $t_k$ . It is not difficult to show (see Chapter 2) that  $p(\underline{x}_k/\underline{z}^k)$  evolves according to the recursion relation

$$p(\underline{x}_k/\underline{z}^k) = \frac{p(\underline{x}_k/\underline{z}^{k-1})p(\underline{z}_k/\underline{x}_k)}{p(\underline{z}_k/\underline{z}^{k-1})} \quad (\text{IV})$$

where

$$p(\underline{x}_k/\underline{z}^{k-1}) = \int p(\underline{x}_k/\underline{z}^{k-1})p(\underline{z}_k/\underline{x}_{k-1}, \underline{u}_{k-1}) d\underline{x}_{k-1}$$

and

$$p(\underline{z}_k/\underline{z}^{k-1}) = \int p(\underline{x}_k/\underline{z}^{k-1})p(\underline{z}_k/\underline{x}_k) d\underline{x}_k$$

The denominator of (IV) does not involve  $\underline{x}_k$  and plays the role of a normalizing constant.

The general concept of dealing with the a posteriori density is referred to as the Bayesian approach to estimation and control. Unfortunately, it is not possible to solve (IV) in closed form for most problems. (The major exception occurs for linear systems.) Furthermore, the computational requirements for solving (IV) numerically become astronomically large for almost any non-trivial problem. Thus, it becomes apparent that approximations must be introduced that will reduce the complexity of the problem without destroying its character.

### 1.3 PREVIEW OF COMING ATTRACTIONS

Since one must know  $p(\underline{x}_k/\underline{z}^k)$  before proceeding with the solution of the estimation and control problems, it is the intention in this investigation to develop a means of approximating the density. It is believed that a combination of perturbative and Bayesian techniques will permit the development of a

theory that is at once more general than the linear theory but more computationally attractive than the general Bayesian approach. The procedure that is proposed for achieving this meld is described below.

- (1) At each sampling time  $t_k$ , nominal values<sup>†</sup> for  $\underline{x}_{k-1}$ ,  $\underline{u}_{k-1}$ , and  $\underline{w}_{k-1}$  are assumed. Then the  $\underline{f}_k$  is expanded in a Taylor Series. The measurement equation  $\underline{h}_k$  is expanded about  $\underline{f}_k$  ( $\underline{x}_{k-1}^*$ ,  $\underline{u}_{k-1}^*$ ,  $\underline{w}_{k-1}^*$ ) and  $\underline{y}_k^*$ .
- (2) A form for  $p(\underline{x}_k/\underline{z}^k)$  must be assumed. This form is required to be true for all  $k$ .
- (3) The a priori statistics for the plant and measurement noise and the expansions of  $\underline{f}_k$  and  $\underline{h}_k$  are introduced into (IV). Only those terms are retained that yield the desired form for  $p(\underline{x}_k/\underline{z}^k)$ .

The application of this procedure to a system leads to several questions concerning the resulting approximation.

- (1) Does the approximation describe  $p(\underline{x}_k/\underline{z}^k)$  accurately enough to have confidence in the validity of the estimation and control policies that are subsequently derived?
- (2) Does the approximation lead to estimation and control policies that provide a significant improvement over linear policies?

A question that is related to the preceding one can be phrased in the following manner.

---

<sup>†</sup> nominal values are denoted by the superscript \*.

- (3) Can special techniques be developed that extend the range of applicability of linear theory and thereby eliminate the need for nonlinear considerations in many problems?

In this study, several specific approximations are developed. Then, these questions are considered by examining the estimates of the state of a dynamical system that are obtained from the approximations. This is accomplished through digital simulation.

In Chapter 2, the general Bayesian approach is discussed. The solution of the minimum mean-square estimation problem is shown to be the conditional mean. Conditions that the control must satisfy for the performance index  $(III_C)$  to be minimized are derived in terms of the a posteriori density. Then, equations which describe the a posteriori density  $p(\underline{x}_k/\underline{z}^{k+\gamma})$  (for any integer  $\gamma$ ) are derived. These results have appeared [33, 14, 32, 35] before in the literature. In addition, the relations describing the  $p(\underline{x}_k/\underline{z}^{k+\gamma})$  are rewritten in terms of characteristic functions. It has been found in Chapter 3 that the characteristic function formulation can reduce the amount of algebraic manipulation required in the solution of a problem.

The Bayesian approach is applied to the linear stochastic control problem in Chapter 3. It is used to obtain the Kalman filter equations [24, 30], Rauch's smoothing equations [36], and to prove the Separation Principle. It is seen from the proof that the Separation Principle is valid because the error covariance matrix for this case does not depend upon the measurement data.

The procedure stated at the start of this section is applied in Chapter 4 under the constraint that  $p(\underline{x}_k/\underline{z}^k)$  is Gaussian. It is demonstrated that the filter equations that are obtained are not the linear Kalman equations. Instead, second order terms appear and the conditional covariance becomes a function of the measurement data. This is a distinct departure from the Kalman filter in which the conditional covariance is independent of the measurements. It is, however, characteristic of nonlinear estimates. It is further observed that a distinct simplification in the filter is obtained by requiring the nominal value for  $\underline{x}_{k-1}$  to be  $\hat{\underline{x}}_{k-1}$ . The control of a linear system with nonlinear measurements is considered, and it is suggested that the Separation Principle is no longer valid.

The problem of estimating the state of a spacecraft moving in a nearly circular, 100 nautical orbit about the Earth from horizon sensor measurements is considered in Chapter 5. A digital computer program simulation was set up to simulate the physical system and the techniques and results obtained in Chapters 3 and 4 are utilized. The linear filter of Chapter 3 and the nonlinear filter of Chapter 4 are compared. In addition, techniques for extending the range of the linear filter are proposed and used. Several interesting conclusions are suggested by these numerical results.

In Chapter 6, attention is restricted to the nonlinear estimation problem. In this chapter, the a posteriori density is approximated by a truncated Edgeworth expansion. All considerations are limited to scalar plant and measurement equations, and approximations retaining third and fourth order

conditional moments are derived. It is seen that the approximation is achieved by developing recurrence relations for the moments of the distribution.

The results of Chapter 6 are applied to a simple problem in Chapter 7. Filters based on a linear theory are exercised and compared with the filters produced by the approximations. "Modified" linear techniques are also examined.

The major results and conclusions provided by this study are summarized in Chapter 8. The contents of each chapter are described in Section 8.1, and the reader might consult that discussion before proceeding through Chapters 2 through 7.

## CHAPTER TWO

## THE BAYESIAN APPROACH

In the so-called "Bayesian approach" to the problems of determining estimation and control policies for stochastic systems, one is concerned first of all with the determination of the a posteriori density function  $p(\underline{x}_k/\underline{z}^{k+\gamma})$ . This density function provides all of the data required for the solution of these problems. To see that this is indeed the case, the following section shall be devoted to the solution of the minimum mean-square estimation problem and the optimal control problem. In this discussion, it is assumed that the necessary density functions are available. The solution of these problems provide a means for determining "best estimates" and "optimal controllers" if the  $p(\underline{x}_k/\underline{z}^{k+\gamma})$  is known. In Section 2.2 it is demonstrated that the a posteriori density can be determined from the a priori statistics specified for the plant and measurement noise. Naturally, the functions  $\underline{f}_k$  and  $\underline{h}_k$  enter these considerations. In the concluding section of this chapter, the relations describing the a posteriori density are rewritten in terms of characteristic functions.

Reference will be made frequently in this and subsequent chapters to three properties of conditional density functions [37].

1. For random variables  $\underline{a}$  and  $\underline{b}$  with joint probability density function  $p(\underline{a}, \underline{b})$ , the conditional density of  $\underline{a}$ , given  $\underline{b}$ , is defined as

$$p(\underline{a}|\underline{b}) = \frac{p(\underline{a}, \underline{b})}{p(\underline{b})} \quad (2.1)$$

2. For random variables  $\underline{a}$ ,  $\underline{b}$ , and  $\underline{c}$ ,

$$p(\underline{a}, \underline{b}|\underline{c}) = p(\underline{b}|\underline{c}) p(\underline{a}|\underline{b}, \underline{c}) \quad (2.2)$$

This is known as the chain rule.

3. For random variables  $\underline{a}$ ,  $\underline{b}$ , and  $\underline{c}$ ,

$$p(\underline{a}|\underline{b}) = \int p(\underline{a}|\underline{b}, \underline{c}) p(\underline{b}|\underline{c}) d\underline{b} \quad (2.3)$$

This is the integrated form of the chain rule and represents one version of the Chapman-Kolmogorov equation.

Note that the definition of conditional densities (2.1) can be rewritten as

$$p(\underline{a}|\underline{b}) = \frac{p(\underline{b}|\underline{a}) p(\underline{a})}{p(\underline{b})} \quad (2.4)$$

This relation is known as Bayes' rule and is the source for the term Bayesian as used in this and other chapters.

Note that the integration indicated in (2.3) involves vector variables.

The single integral sign will be used for both scalar and vector variables and  $d\underline{b}$  will be used to describe the differential  $db_1 db_2 \dots db_n$ . When more than one vector is involved, the differential will be written as  $d(\underline{a}, \underline{b}, \dots, \underline{z})$ .

## 2.1 OPTIMAL ESTIMATION AND CONTROL FOR STOCHASTIC TIME-DISCRETE SYSTEMS

The mean-square estimation problem and the optimal control problem shall be solved in this section in terms of the a posteriori density function.

### 2.1.1 The Minimum Mean-Square Estimation Problem

The solution of the minimum mean-square estimation problem [38] is provided by the following lemma.



LEMMA 2.1: Suppose that a random variable  $\underline{x}$  is to be estimated from the known variables\*  $\underline{z}^q$ . The  $\underline{x}$  and  $\underline{z}^q$  have the joint probability density function  $p(\underline{x}, \underline{z}^q)$ . The estimate  $\hat{\underline{x}}$  is to be chosen as a function of the  $\underline{z}^q$  so that

$$E[(\hat{\underline{x}} - \underline{x})^T (\hat{\underline{x}} - \underline{x})] = \text{minimum}$$

Then, the mean-square estimate of  $\underline{x}$  is

$$\hat{\underline{x}} = E[\underline{x} | \underline{z}^q] \quad (V_E)$$

Proof: Write  $E[(\hat{\underline{x}} - \underline{x})^T (\hat{\underline{x}} - \underline{x})]$  in terms of the joint density function.

$$E[(\hat{\underline{x}} - \underline{x})^T (\hat{\underline{x}} - \underline{x})] = \int (\hat{\underline{x}} - \underline{x})^T (\hat{\underline{x}} - \underline{x}) p(\underline{x}, \underline{z}^q) d(\underline{x}, \underline{z}^q) \quad (2.5)$$

From (2.1), the density function can be written as

$$p(\underline{x}, \underline{z}^q) = P(\underline{x} | \underline{z}^q) p(\underline{z}^q)$$

Thus, (2.5) is equivalent to

$$E[(\hat{\underline{x}} - \underline{x})^T (\hat{\underline{x}} - \underline{x})] = \int \{ \int (\hat{\underline{x}} - \underline{x})^T (\hat{\underline{x}} - \underline{x}) p(\underline{x} | \underline{z}^q) d\underline{x} \} p(\underline{z}^q) d\underline{z}^q$$

Consider the integral in brackets.. Since  $\hat{\underline{x}}$  depends only upon the  $\underline{z}^q$ ,

the integral can be written as

$$\begin{aligned} \int (\hat{\underline{x}} - \underline{x})^T (\hat{\underline{x}} - \underline{x}) p(\underline{x} | \underline{z}^q) d\underline{x} &= \hat{\underline{x}}^T \hat{\underline{x}} - 2\hat{\underline{x}}^T E[\underline{x} | \underline{z}^q] + E[\underline{x}^T \underline{x} | \underline{z}^q] \\ &= (\hat{\underline{x}} - E[\underline{x} | \underline{z}^q])^T (\hat{\underline{x}} - E[\underline{x} | \underline{z}^q]) + E[\underline{x}^T \underline{x} | \underline{z}^q] \\ &\quad - \{E[\underline{x} | \underline{z}^q]\}^T E[\underline{x} | \underline{z}^q] \end{aligned} \quad (2.6)$$

By definition this quantity is positive, so to minimize  $E[(\hat{\underline{x}} - \underline{x})^T (\hat{\underline{x}} - \underline{x})]$ , it is sufficient to minimize (2.6). Only the first term involves  $\hat{\underline{x}}$ , and the smallest value that it can assume is zero. Thus, the minimizing estimate is given by

$$\hat{\underline{x}} = E[\underline{x} | \underline{z}^q] \quad \text{Q. E. D.}$$

---

\* Recall that the set  $(\underline{z}_1, \underline{z}_2, \dots, \underline{z}_q)$  is denoted by  $\underline{z}^q$ .

This lemma shows that the conditional mean provides the mean-square estimate of  $\underline{x}$ . Certainly, if one knows the a posteriori density function  $p(\underline{x}/\underline{z}^q)$ , then the estimation problem has in principle been solved. Other estimates such as those given by the mode or by the median of the distribution are also established from knowledge of  $p(\underline{x}/\underline{z}^q)$ .

The conditional mean provides an unbiased estimate of a variable  $\underline{x}$ . That is, it is true that

$$E[\underline{x}] = E[\hat{\underline{x}}].$$

This is verified in the following manner.

By definition, one has

$$E[\hat{\underline{x}}] = \int \hat{\underline{x}} p(\underline{z}^q) d\underline{z}^q$$

But  $\hat{\underline{x}}$  is the conditional mean, so

$$E[\hat{\underline{x}}] = \int \left\{ \int \underline{x} p(\underline{x}/\underline{z}^q) d\underline{x} \right\} p(\underline{z}^q) d\underline{z}^q$$

From 2.1, it follows that

$$E[\hat{\underline{x}}] = \int \underline{x} p(\underline{x}, \underline{z}^q) d(\underline{x}, \underline{z}^q)$$

Integrate with respect to  $\underline{z}^q$ . Then,

$$E[\hat{\underline{x}}] = \int \underline{x} p(\underline{x}) d\underline{x}$$

$$\stackrel{\text{Df}}{=} E[\underline{x}]$$

### 2.1.2 The Control Problem

Suppose that a feedback control policy  $\underline{u}^{N-1}$  is to be determined for the system (I) and (II) that minimizes  $(III)_C$ . It shall be shown that

LEMMA 2.2: The optimal feedback control policy for the system (I) - (II) and performance index  $(III)_C$  is the control that causes

$$E[\bar{J}_{N-k+1}/\underline{z}^{N-k}] = \text{minimum} \quad (V_C)$$

where

$$\bar{J}_{N-k+1} \stackrel{\text{Df}}{=} \int (w_{N-k+1} + \bar{J}_{N-k+1}) \delta(\underline{x}_{N-k+1} - \underline{f}_{N-k+1}) p(\underline{w}_{N-k}) d(\underline{x}_{N-k+1}, \underline{w}_{N-k})$$

and

$$\bar{J}_{N-k+1}^o \stackrel{\text{Df}}{=} \int \bar{J}_{N-k+2}^o \delta(\underline{z}_{N-k+1} - \underline{h}_{N-k+1}) p(\underline{v}_{N-k+1}) d(\underline{v}_{N-k+1}, \underline{z}_{N-k+1})$$

At the last stage, the  $\bar{J}_N$  is defined to be zero. The  $\delta(\cdot)$  represents the Dirac delta function. The superscript o on  $\bar{J}_{N-k+2}^o$  is used to signify that it has been evaluated with the optimal control  $\underline{u}_{N-k+1}^o$ . The cost associated with the optimal control is

$$E[V_{k-1}^o] = E\{E[\bar{J}_{N-k+2}^o/\underline{z}^{N-k+1}]\}$$

**Proof:** This assertion is proved inductively. Consider the control for the last stage. First, from (III<sub>C</sub>)

$$\begin{aligned} E[V_N] &= E\left[\sum_{i=1}^N W_i(\underline{x}_i, \underline{u}_{i-1})\right] \\ &= \int \left[\sum_{i=1}^N W_i(\underline{x}_i, \underline{u}_{i-1})\right] p(\underline{x}^N, \underline{z}^{N-1}) d(\underline{x}^N, \underline{z}^{N-1}) \\ &= \int \left[\sum_{i=1}^{N-1} W_i(\underline{x}_i, \underline{u}_{i-1})\right] p(\underline{x}^N, \underline{z}^{N-1}) d(\underline{x}^N, \underline{z}^{N-1}) \\ &\quad + \int W_N(\underline{x}_N, \underline{u}_{N-1}) p(\underline{x}^N, \underline{z}^{N-1}) d(\underline{x}^N, \underline{z}^{N-1}) \end{aligned} \quad (2.7)$$

The integrand of the first term does not contain  $\underline{x}_N$  or  $\underline{z}^{N-1}$ , so an integration will eliminate these variables. The control  $\underline{u}_{N-1}$  enters only the last integral so no other terms need to be considered in determining the optimal control for the last stage. Let

$$E[V_1] \stackrel{\text{Df}}{=} \int W_N(\underline{x}_N, \underline{u}_{N-1}) p(\underline{x}^N, \underline{z}^{N-1}) d(\underline{x}^N, \underline{z}^{N-1}) \quad (2.8)$$

Using the integrated chain rule (2.3),  $p(\underline{x}^N, \underline{z}^{N-1})$  can be written as

$$\begin{aligned} p(\underline{x}^N, \underline{z}^{N-1}) &= p(\underline{x}_N / \underline{x}^{N-1}, \underline{z}^{N-1}) p(\underline{x}^{N-1}, \underline{z}^{N-1}) \\ &= p(\underline{x}^{N-1}, \underline{z}^{N-1}) \int p(\underline{x}_N / \underline{x}^{N-1}, \underline{z}^{N-1}, \underline{w}_{N-1}) \\ &\quad p(\underline{w}_{N-1} / \underline{x}^{N-1}, \underline{z}^{N-1}) d\underline{w}_{N-1} \end{aligned}$$

But  $\underline{w}_{N-1}$  is a white noise sequence, so

$$p(\underline{w}_{N-1} / \underline{x}^{N-1}, \underline{z}^{N-1}) = p(\underline{w}_{N-1})$$

and since  $\underline{z}^{N-1}$  defines  $\underline{u}_{N-1}$ , it is clear from (I) that

$$\begin{aligned} p(\underline{x}_N / \underline{x}^{N-1}, \underline{z}^{N-1}, \underline{w}_{N-1}) &= p(\underline{x}_N / \underline{x}_{N-1}, \underline{u}_{N-1}, \underline{w}_{N-1}) \\ &= \delta(\underline{x}_N - \underline{f}_N) \end{aligned}$$

where  $\delta(\cdot)$  represents the Dirac delta function. Thus,

$$p(\underline{x}^N, \underline{z}^{N-1}) = p(\underline{x}^{N-1}, \underline{z}^{N-1}) \int \delta(\underline{x}_N - \underline{f}_N) p(\underline{w}_{N-1}) d\underline{w}_{N-1}$$

Substituting this into (2.8),  $E[V_1]$  becomes

$$E[V_1] = \int W_N(\underline{x}_N, \underline{u}_{N-1}) \delta(\underline{x}_N - \underline{f}_N) p(\underline{w}_{N-1}) p(\underline{x}^{N-1}, \underline{z}^{N-1}) d(\underline{x}^N, \underline{z}^{N-1}, \underline{w}_{N-1})$$

Let

$$\bar{z}_N \stackrel{\text{Df}}{=} \int W_N(\underline{x}_N, \underline{u}_{N-1}) \delta(\underline{x}_N - \underline{f}_N) p(\underline{w}_{N-1}) d(\underline{x}_N, \underline{w}_{N-1})$$

Then

$$E[V_1] = \int \bar{z}_N p(\underline{x}_{N-1}, \underline{z}^{N-1}) d(\underline{x}_{N-1}, \underline{z}^{N-1})$$

where the integration with respect to  $\underline{x}^{N-2}$  has been performed. This can be further modified to

$$E[V_1] = \int \left\{ \int \bar{z}_N p(\underline{x}_{N-1} / \underline{z}^{N-1}) d\underline{x}_{N-1} \right\} p(\underline{z}^{N-1}) d\underline{z}^{N-1} \quad (2.9)$$

The  $E[V_1]$  will be minimized by the control  $\underline{u}_{N-1}^0$  that causes

$$E[\bar{z}_N / \underline{z}^{N-1}] = \text{minimum}$$

This verifies  $(V_C)$  for  $k = 1$ . Denote the value of  $\bar{z}_N$  that is evaluated with  $\underline{u}_{N-1}^0$  by  $\bar{z}_N^0$  so that

$$E[V_1^0] = E\{E[\bar{z}_N^0 / \underline{z}^{N-1}]\}$$

Suppose that the optimal controls  $\underline{u}_{N-k+1}^0, \dots, \underline{u}_{N-1}^0$  are computed according to  $(V_C)$  and that the expected cost associated with these  $(k-1)$  stages is

$$E[V_{k-1}^0] = E\{E[\bar{z}_{N-k+2}^0 / \underline{z}^{N-k+1}]\}$$

Then, using the Principle of Optimality, it follows that

$$\begin{aligned} E[V_k] &= \int W_{N-k+1}(\underline{x}_{N-k+1}, \underline{u}_{N-k}) p(\underline{x}^{N-k+1}, \underline{z}^{N-k}) d(\underline{x}^{N-k+1}, \underline{z}^{N-k}) \\ &\quad + E[V_{k-1}^0] \end{aligned} \quad (2.10)$$

Let us rewrite the second term

$$E[V_{k-1}^0] = \int \bar{z}_{N-k+2}^0 p(\underline{x}^{N-k+1}, \underline{z}^{N-k+1}) d(\underline{x}^{N-k+1}, \underline{z}^{N-k+1})$$

But

$$\begin{aligned}
 p(\underline{x}^{N-k+1}, \underline{z}^{N-k+1}) &= p(\underline{z}_{N-k+1} | \underline{x}^{N-k+1}, \underline{z}^{N-k}) p(\underline{x}^{N-k+1}, \underline{z}^{N-k}) \\
 &= p(\underline{x}^{N-k+1}, \underline{z}^{N-k}) \int p(\underline{z}_{N-k+1} / \underline{x}^{N-k+1}, \underline{z}^{N-k}, \underline{v}_{N-k+1}) \\
 &\quad p(\underline{v}_{N-k+1} / \underline{x}^{N-k+1}, \underline{z}^{N-k}) d\underline{v}_{N-k+1}
 \end{aligned}$$

From the assumption on the noise and from (II), this reduces to

$$\begin{aligned}
 p(\underline{x}^{N-k+1}, \underline{z}^{N-k+1}) &= p(\underline{x}^{N-k+1}, \underline{z}^{N-k}) \int \delta(\underline{z}_{N-k+1} - \underline{h}_{N-k+1}) \\
 &\quad p(\underline{v}_{N-k+1}) d\underline{v}_{N-k+1}
 \end{aligned}$$

Using this result,  $E[V_{k-1}^0]$  becomes

$$\begin{aligned}
 E[V_{k-1}^0] &= \int \underline{z}_{N-k+2}^0 \delta(\underline{z}_{N-k+1} - \underline{h}_{N-k+1}) p(\underline{v}_{N-k+1}) p(\underline{x}^{N-k+1} / \underline{z}^{N-k}) \\
 &\quad d(\underline{v}_{N-k+1}, \underline{x}^{N-k+1}, \underline{z}^{N-k+1}).
 \end{aligned}$$

Let

$$\underline{z}_{N-k+1}^0 \stackrel{\text{Df}}{=} \int \underline{z}_{N-k+2}^0 \delta(\underline{z}_{N-k+1} - \underline{h}_{N-k+1}) p(\underline{v}_{N-k+1}) d(\underline{v}_{N-k+1}, \underline{z}_{N-k+1})$$

So

$$E[V_{k-1}^0] = \int \underline{z}_{N-k+1}^0 p(\underline{x}^{N-k+1}, \underline{z}^{N-k}) d(\underline{x}^{N-k+1}, \underline{z}^{N-k}) \quad (2.11)$$

Introducing (2.11) into (2.10) yields

$$E[V_k] = \int (\underline{w}_{N-k+1} + \underline{z}_{N-k+1}^0) p(\underline{x}^{N-k+1}, \underline{z}^{N-k}) d(\underline{x}^{N-k+1}, \underline{z}^{N-k})$$

Proceed as was done to obtain (2.9). It follows that

$$p(\underline{x}^{N-k+1}, \underline{z}^{N-k}) = p(\underline{x}^{N-k}, \underline{z}^{N-k}) \int \delta(\underline{x}_{N-k+1} - \underline{f}_{N-k+1}) p(\underline{w}_{N-k}) d\underline{w}_{N-k} \quad (2.12)$$

Let

$$\begin{aligned} \bar{z}_{N-k+1} &\stackrel{\text{Df}}{=} \int (w_{N-k+1} + \bar{z}_{N-k+1}) \delta(x_{N-k+1} - f_{N-k+1}) \\ &\quad p(w_{N-k}, x_{N-k+1}) \end{aligned}$$

so that

$$\begin{aligned} E[V_k] &= \int \bar{z}_{N-k+1} p(x_{N-k}, \underline{z}^{N-k}) d(x_{N-k}, \underline{z}^{N-k}) \\ &= \int \left\{ \int \bar{z}_{N-k+1} p(x_{N-k} / \underline{z}^{N-k}) dx_{N-k} \right\} p(\underline{z}^{N-k}) d\underline{z}^{N-k} \end{aligned}$$

The optimal control must satisfy  $(V_C)$  and the cost is

$$E[V_k^*] = E\{E[\bar{z}_{N-k+1}^0 / \underline{z}^{N-k}]\}$$

This completes the proof of the lemma.

Q. E. D.

The optimal feedback control problem has been solved in principle if the a posteriori density  $p(x_k / \underline{z}^k)$  is known for all  $k$ . Similar results can be found in [5, 13, 33].

## 2.2 THE A POSTERIORI CONDITIONAL DENSITY FUNCTION

In the preceding section, it was shown that the a posteriori density function provides all of the information required to determine optimal estimation and control policies. In this section, equations governing the structure of  $p(x_k / \underline{z}^{k+y})$  shall be derived.

### 2.2.1 Recursion Relation for $p(x_k / \underline{z}^k)$

The density  $p(x_k / \underline{z}^k)$  can be described by an integral recurrence relation.

This fact shall be stated as a lemma and then proven.

LEMMA 2.3: For the system (I) - (II) and a nonrandomized control policy, the a posteriori density function  $p(\underline{x}_k/\underline{z}^k)$  evolves according to

$$p(\underline{x}_k/\underline{z}^k) = \frac{p(\underline{x}_k/\underline{z}^{k-1})p(\underline{z}_k/\underline{x}_k)}{p(\underline{z}_k/\underline{z}^{k-1})} \quad (\text{IV})$$

where

$$p(\underline{x}_k/\underline{z}^{k-1}) = \int p(\underline{x}_k/\underline{x}_{k-1}, \underline{u}_{k-1})p(\underline{x}_{k-1}/\underline{z}^{k-1})d\underline{x}_{k-1} \quad (2.13)$$

and

$$p(\underline{z}_k/\underline{z}^{k-1}) = \int p(\underline{x}_k/\underline{z}^{k-1})p(\underline{z}_k/\underline{x}_k)d\underline{x}_k \quad (2.14)$$

The initial condition  $p(\underline{x}_0/\underline{z}_0)$  is given by

$$p(\underline{x}_0/\underline{z}_0) = \frac{p(\underline{z}_0/\underline{x}_0)p(\underline{x}_0)}{p(\underline{z}_0)} \quad (2.15)$$

where

$$p(\underline{z}_0) = \int p(\underline{z}_0/\underline{x}_0)p(\underline{x}_0)d\underline{x}_0$$

Proof: The initial condition can be established directly from Bayes rule (2.4).

Thus, consider arbitrary  $k$ .

From the chain rule (2.2), one sees that

$$p(\underline{x}_k, \underline{z}_k/\underline{z}^{k-1}) = p(\underline{x}_k/\underline{z}^k)p(\underline{z}_k/\underline{z}^{k+1})$$

so

$$p(\underline{x}_k/\underline{z}^k) = \frac{p(\underline{x}_k, \underline{z}_k/\underline{z}^{k-1})}{p(\underline{z}_k/\underline{z}^{k-1})}$$

But the chain rule also enables us to write



$$p(\underline{x}_k, \underline{z}_k / \underline{z}^{k-1}) = p(\underline{z}_k / \underline{x}_k, \underline{z}^{k-1}) p(\underline{x}_k / \underline{z}^{k-1})$$

This can be simplified to

$$p(\underline{x}_k, \underline{z}_k / \underline{z}^{k-1}) = p(\underline{z}_k / \underline{x}_k) p(\underline{x}_k / \underline{z}^{k-1})$$

since, from the noise assumptions in (I) and (II), it is true that  $\underline{z}_k$  given  $\underline{x}_k$  is independent of  $\underline{z}^{k-1}$ . Thus,

$$p(\underline{x}_k / \underline{z}^k) = \frac{p(\underline{x}_k / \underline{z}^{k-1}) p(\underline{z}_k / \underline{x}_k)}{p(\underline{z}_k / \underline{z}^{k-1})} \quad (\text{IV})$$

This relation proves (IV). It remains to verify (2.13) and (2.14).

From the integrated chain rule (2.3),

$$p(\underline{x}_k / \underline{z}^{k-1}) = \int p(\underline{x}_k / \underline{x}_{k-1}, \underline{z}^{k-1}) p(\underline{x}_{k-1} / \underline{z}^{k-1}) d\underline{x}_{k-1}$$

But  $\underline{z}^{k-1}$  defines  $\underline{u}_{k-1}$  so

$$p(\underline{x}_k / \underline{z}^{k-1}) = \int p(\underline{x}_k / \underline{x}_{k-1}, \underline{u}_{k-1}) p(\underline{x}_{k-1} / \underline{z}^{k-1}) d\underline{x}_{k-1}$$

The integrated chain rule also allows one to write

$$p(\underline{z}_k / \underline{z}^{k-1}) = \int p(\underline{x}_k / \underline{z}^{k-1}) p(\underline{z}_k / \underline{x}_k) d\underline{x}_k$$

This completes the proof.

Q. E. D.

The  $p(\underline{z}_k / \underline{z}^{k-1})$  in (IV) does not depend upon  $\underline{x}_k$ , so it can be seen to be nothing more than a normalization constant. The basic structure for the recursion relation is provided by the numerator. It should be noted in passing that it is not possible in general to perform the integration indicated in (2.13) to obtain a closed-form for  $p(\underline{x}_k / \underline{z}^{k-1})$ .

### 2.2.2 The A Posteriori Density Function for Prediction and Smoothing

Equation (IV) provides the basic formula for filtering and control purposes. Occasions do arise when it is desirable to obtain predicted or smoothed estimates of  $\underline{x}_k$ , so it is necessary to determine the density  $p(\underline{x}_k/\underline{z}^{k+\gamma})$  for  $\gamma \neq 0$ . For this case the control variables will be eliminated thereby reducing the plant equation to

$$\underline{x}_k = \underline{f}_k(\underline{x}_{k-1}, \underline{w}_{k-1}) \quad (I_E)$$

LEMMA 2.4: For the system  $(I_E) - (II)$ , the a posteriori density function

$p(\underline{x}_k/\underline{z}^{k-\gamma})$  for  $\gamma > 0$  is

$$p(\underline{x}_k/\underline{z}^{k-\gamma}) = \int p(\underline{x}_k/\underline{x}_{k-1})p(\underline{x}_{k-1}/\underline{x}_{k-2}) \cdots p(\underline{x}_{k-\gamma+1}/\underline{x}_{k-\gamma}) \\ p(\underline{x}_{k-\gamma}/\underline{z}^{k-\gamma})d(\underline{x}_{k-1}, \dots, \underline{x}_{k-\gamma}) \quad (VI_E)$$

The proof of this statement follows immediately from the repeated application of the integrated chain rule. See the derivation of (2.13) for the case when  $\gamma = 1$ .

The derivation of the smoothing density is somewhat more involved. The result can be stated as follows.

LEMMA 2.5: For the system  $(I_E) - (II)$  the a posteriori conditional density function  $p(\underline{x}_{k-\gamma}/\underline{z}^k)$  for  $\gamma > 0$  is given by

$$p(\underline{x}_{k-\gamma}/\underline{z}^k) = \frac{p(\underline{x}_{k-\gamma}/\underline{z}^{k-\gamma})p(\underline{z}_k, \dots, \underline{z}_{k-\gamma+1}/\underline{x}_{k-\gamma})}{p(\underline{z}_k/\underline{z}^{k-1}), \dots, p(\underline{z}_{k-\gamma+1}/\underline{z}^{k-\gamma})} \quad (VII_E)$$

where  $p(\underline{z}_k, \dots, \underline{z}_{k-\gamma+1}/\underline{x}_{k-\gamma})$  is computed recursively according to

$$p(z_k, \dots, z_{k-\gamma+1}/x_{k-\gamma}) = \int p(z_k, \dots, z_{k-\gamma+2}/x_{k-\gamma+1}) p(z_{k-\gamma+1}/x_{k-\gamma+1}) \\ p(x_{k-\gamma+1}/x_{k-\gamma}) dx_{k-\gamma+1} \quad (2.16)$$

The initial condition for this relation is (i.e.,  $\gamma = 1$ )

$$p(z_k/x_{k-1}) = \int p(z_k/x_k) p(x_k/x_{k-1}) dx_k \quad (2.17)$$

**Proof:** The proof shall be inductive. Let  $\gamma = 1$  and consider

$$p(x_{k-1}, x_k, z^k) = p(x_{k-1}, x_k, z_k/z^{k-1}) p(z^{k-1}) \\ = p(x_k, z_k/x_{k-1}, z^{k-1}) p(x_{k-1}/z^{k-1}) p(z^{k-1}) \\ = p(x_k, z_k/x_{k-1}) p(x_{k-1}/z^{k-1}) p(z^{k-1}) \\ = p(z_k/x_k, x_{k-1}) p(x_k/x_{k-1}) p(x_{k-1}/z^{k-1}) p(z^{k-1}) \\ = p(z_k/x_k) p(x_k/x_{k-1}) p(x_{k-1}/z^{k-1}) p(z^{k-1}) \quad (2.18)$$

Furthermore, it is true that

$$p(x_{k-1}, x_k, z^k) = p(x_{k-1}, x_k/z^k) p(z^k) \\ = p(x_{k-1}, x_k/z^k) p(z_k/z^{k-1}) p(z^{k-1}) \quad (2.19)$$

Equating (2.18) and (2.19) and rearranging terms, one obtains

$$p(x_{k-1}, x_k/z^k) = \frac{p(z_k/x_k) p(x_k/x_{k-1}) p(x_{k-1}/z^{k-1})}{p(z_k/z^{k-1})}$$

Integrate with respect to  $x_k$ . Then

$$p(x_{k-1}/z^k) = \frac{p(x_{k-1}/z^{k-1}) p(z_k/x_{k-1})}{p(z_k/z^{k-1})}$$

where, by the integrated chain rule (2.3),

$$p(\underline{z}_k / \underline{x}_{k-1}) = \int p(\underline{z}_k / \underline{x}_k) p(\underline{x}_k / \underline{x}_{k-1}) d\underline{x}_k$$

This proves (2.16) and (VII<sub>E</sub>) for  $\gamma = 1$ .

Suppose that (VII<sub>E</sub>) and (2.16) are true for  $\gamma = j-1$  and let  $\gamma = j$ . Proceed as for  $\gamma = 1$ . Then it follows that

$$\begin{aligned} p(\underline{x}_{k-j}, \underline{x}_{k-j+1}, \underline{z}^k) &= p(\underline{z}_k, \dots, \underline{z}_{k-j+1} / \underline{x}_{k-j+1}) \\ &\quad p(\underline{x}_{k-j+1} / \underline{x}_{k-j}) p(\underline{x}_{k-j} / \underline{z}^{k-j}) p(\underline{z}^{k-j}) \end{aligned} \quad (2.20)$$

and, also that

$$\begin{aligned} p(\underline{x}_{k-j}, \underline{x}_{k-j+1}, \underline{z}^k) &= p(\underline{x}_{k-j}, \underline{x}_{k-j+1} / \underline{z}^k) p(\underline{z}_k / \underline{z}^{k-1}) \dots \\ &\quad p(\underline{z}_{k-j+1} / \underline{z}^{k-j}) p(\underline{z}^{k-j}) \end{aligned} \quad (2.21)$$

Equate (2.20) and (2.21) to obtain

$$p(\underline{x}_{k-j}, \underline{x}_{k-j+1} / \underline{z}^k) = \frac{p(\underline{x}_{k-j} / \underline{z}^{k-j}) p(\underline{x}_{k-j+1} / \underline{x}_{k-j}) p(\underline{z}_k, \dots, \underline{z}_{k-j+1} / \underline{x}_{k-j+1})}{p(\underline{z}_k / \underline{z}^{k-j}) \dots p(\underline{z}_{k-j+1} / \underline{z}^{k-j})}$$

Integrate with respect to  $\underline{x}_{k-j+1}$ . Then

$$p(\underline{x}_{k-j} / \underline{z}^k) = \frac{p(\underline{x}_{k-j} / \underline{z}^{k-j}) p(\underline{z}_k, \dots, \underline{z}_{k-j+1} / \underline{x}_{k-j})}{p(\underline{z}_k / \underline{z}^{k-1}) \dots p(\underline{z}_{k-j+1} / \underline{z}^{k-j})}$$

where from the integrated chain rule, it is true that

$$\begin{aligned} p(\underline{z}_k, \dots, \underline{z}_{k-j+1} / \underline{x}_{k-j}) &= \int p(\underline{z}_k, \dots, \underline{z}_{k-j+1} / \underline{x}_{k-j+1}) \\ &\quad p(\underline{x}_{k-j+1} / \underline{x}_{k-j}) d\underline{x}_{k-j+1} \end{aligned}$$

By the chain rule (2.2), it is apparent that

$$p(\underline{z}_k, \dots, \underline{z}_{k-j+1} / \underline{x}_{k-j+1}) = p(\underline{z}_k, \dots, \underline{z}_{k-j+2} / \underline{x}_{k-j+1}) p(\underline{z}_{k-j+1} / \underline{x}_{k-j+1})$$

This completes the proof of VII<sub>E</sub> and (2.16).

Q. E. D.

### 2.3 CHARACTERISTIC FUNCTION EQUIVALENTS

Relations for the a posteriori conditional density function  $p(\underline{x}_k / \underline{z}^{k+\gamma})$  were derived in the preceding section. It is, of course, possible to obtain from these relations their characteristic function equivalents. These relations are to be derived and exhibited in this section. The characteristic functions are introduced primarily for future reference. It has been found that in many cases the problem solutions are most easily obtained using the characteristic function formulation. The reader is encouraged to perform the derivations in Chapter 3 by using the probability density relations of Section 2.2.

Recall that the characteristic function  $\varphi$  and the probability density function  $p$  associated with a random variable  $\underline{x}$  form a Fourier transform pair [10].

$$\left. \begin{aligned} \varphi(\underline{s}) &\stackrel{\text{Df}}{=} E[\exp(i \underline{s}^T \underline{x})] = \int_{-\infty}^{\infty} \exp(i \underline{s}^T \underline{x}) p(\underline{x}) d\underline{x} \\ \text{and} \\ p(\underline{x}) &= \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} \exp(-i \underline{s}^T \underline{x}) \varphi(\underline{s}) d\underline{s} \end{aligned} \right\} \quad (2.22)$$

Consider the characteristic function for  $p(\underline{x}_k / \underline{z}^k)$  as described by (IV).

LEMMA 2.6: The characteristic function  $\varphi(\underline{s}_k)$  for  $p(\underline{x}_k / \underline{z}^k)$  is

$$\begin{aligned} \varphi(\underline{s}_k) &= \frac{1}{(2\pi)^{n+m} p(\underline{z}_k / \underline{z}^{k-1})} \int \exp[-i(\underline{s}_{k/k-1} - \underline{s}_k)^T \underline{x}_k - i \underline{s}_v^T \underline{z}_k] \\ &\quad \varphi(\underline{s}_{k/k-1}) \varphi(\underline{s}_v) d(\underline{x}_k, \underline{s}_v, \underline{s}_{k/k-1}) \end{aligned} \quad (\text{VIII})$$

where  $\varphi(\underline{s}_{k/k-1})$  and  $\varphi(\underline{s}_v)$  are the characteristic functions associated with  $p(\underline{x}_k/\underline{z}^{k-1})$  and  $p(\underline{z}_k/\underline{x}_k)$ , respectively. The characteristic function  $\varphi(\underline{s}_{k/k-1})$  is given by

$$\begin{aligned} \varphi(\underline{s}_{k/k-1}) &= \frac{1}{(2\pi)^{2n}} \int \exp[-i(\underline{s}_w - \underline{s}_{k/k-1})^T \underline{x}_k - i \underline{s}_{k-1}^T \underline{x}_{k-1}] \\ &\quad \varphi(\underline{s}_w) \varphi(\underline{s}_{k-1}) d(\underline{x}_{k-1}, \underline{x}_k, \underline{s}_w, \underline{s}_{k-1}) \end{aligned} \quad (2.23)$$

where  $\varphi(\underline{s}_w)$  is the characteristic function associated with  $p(\underline{x}_k/\underline{x}_{k-1}, \underline{u}_{k-1})$ .

Also,

$$\begin{aligned} p(\underline{z}_k/\underline{z}^{k-1}) &= \frac{1}{(2\pi)^{n+m}} \int \exp[-i \underline{s}_{k/k-1}^T \underline{x}_k - i \underline{s}_v^T \underline{z}_k] \\ &\quad \varphi(\underline{s}_{k/k-1}) \varphi(\underline{s}_v) d(\underline{x}_k, \underline{s}_v, \underline{s}_{k/k-1}) \end{aligned} \quad (2.24)$$

The characteristic function  $\varphi(\underline{s}_o)$  for  $p(\underline{x}_o/\underline{z}_o)$  is

$$\begin{aligned} \varphi(\underline{s}_o) &= \frac{1}{(2\pi)^{n+m} p(\underline{z}_o)} \int \exp[-i(\underline{s}_m - \underline{s}_o)^T \underline{x}_o - i \underline{s}_v^T \underline{z}_o] \\ &\quad \varphi(\underline{s}_m) \varphi(\underline{s}_v) d(\underline{s}_m, \underline{s}_v) \end{aligned} \quad (2.25)$$

The  $\varphi(\underline{s}_m)$  is the characteristic function for  $p(\underline{x}_o)$ .

**Proof:** The proof follows directly from the definitions (2.22) and from (IV).

The characteristic function of  $p(\underline{x}_k/\underline{z}^k)$  is

$$\begin{aligned} \varphi(\underline{s}_k) &= \int \exp[i \underline{s}_k^T \underline{x}_k] p(\underline{x}_k/\underline{z}^k) d\underline{x}_k \\ &= \frac{1}{p(\underline{z}_k/\underline{z}^{k-1})} \int \exp[i \underline{s}_k^T \underline{x}_k] p(\underline{x}_k/\underline{z}^{k-1}) p(\underline{z}_k/\underline{x}_k) d\underline{x}_k \end{aligned} \quad (2.26)$$

But

$$p(\underline{x}_k/\underline{z}^{k-1}) = \frac{1}{(2\pi)^n} \int \exp[-i \underline{s}_{k/k-1}^T \underline{x}_k] \varphi(\underline{s}_{k/k-1}) d\underline{s}_{k/k-1} \quad (2.27)$$

and from (II) it is clear that

$$p(\underline{z}_k/\underline{x}_k) = \frac{1}{(2\pi)^m} \int \exp[-i \underline{s}_v^T \underline{z}_k] \varphi(\underline{s}_v) d\underline{s}_v \quad (2.28)$$

Substitute (2.27) and (2.28) into (2.26) and (VIII) follows directly.

The characteristic function  $\varphi(\underline{s}_{k/k-1})$  is given by

$$\varphi(\underline{s}_{k/k-1}) = \int \exp[i \underline{s}_{k/k-1}^T \underline{x}_k] p(\underline{x}_k/\underline{z}^{k-1}) d\underline{x}_k$$

But from (I)

$$p(\underline{x}_k/\underline{x}_{k-1}, \underline{u}_{k-1}) = \frac{1}{(2\pi)^n} \int \exp[-i \underline{s}_w^T \underline{x}_k] \varphi(\underline{s}_w) d\underline{s}_w$$

so from (2.13)

$$\begin{aligned} \varphi(\underline{s}_{k/k-1}) &= \frac{1}{(2\pi)^{2n}} \int \exp[-i (\underline{s}_w - \underline{s}_{k/k-1})^T \underline{x}_k - i \underline{s}_{k-1}^T \underline{x}_{k-1}] \\ &\quad \varphi(\underline{s}_w) \varphi(\underline{s}_{k-1}) d(\underline{x}_k, \underline{x}_{k-1}, \underline{s}_w, \underline{s}_{k-1}) \end{aligned}$$

This proves (2.23). The  $p(\underline{z}_k/\underline{z}^{k-1})$  can be written in terms of characteristic functions directly from (2.14). The characteristic function  $\varphi(\underline{s}_0)$  follows immediately from (2.15) Q. E. D.

The characteristic function for the smoothing density follows immediately from (VI).

**LEMMA 2.7:** The characteristic function equivalent of  $p(\underline{x}_k/\underline{z}^{k-\gamma})$  for  $\gamma > 0$  is

$$\begin{aligned} \varphi(\underline{s}_{k/k-\gamma}) &= \frac{1}{(2\pi)^{\gamma n}} \int \exp[-i (\underline{s}_{wk-1} - \underline{s}_k)^T \underline{x}_k - i \underline{s}_{wk-2}^T \underline{x}_{k-1}] \dots \\ &\quad \exp[-i \underline{s}_{wk-\gamma}^T \underline{x}_{k-\gamma+1} - i \underline{s}_w^T \underline{x}_{k-\gamma}] \\ &\quad d(\underline{x}_{k-1}, \dots, \underline{x}_{k-\gamma}, \underline{s}_{wk-1}, \dots, \underline{s}_{k/k-\gamma}) \end{aligned} \quad (IX_E)$$

LEMMA 2.8: The characteristic function for the smoothing density is

$$\begin{aligned} \varphi(\underline{s}_{k-\gamma}/k) &= \frac{1}{(2\pi)^{[(\gamma+1)m+n]} \prod_{j=k-\gamma}^{k-1} p(\underline{z}_{j+1}/\underline{z}_j)} \int \exp[-i(\underline{s}_{k-\gamma} - \underline{s}_{k-\gamma}/k)^T \underline{x}_{k-\gamma}] \\ &\quad \exp[-i \sum_{j=k-\gamma}^{k-1} \underline{\sigma}_{j+1/j}^T \underline{z}_{j+1}] \varphi(\underline{s}_{k-\gamma}) \varphi(\underline{\sigma}_{k-\gamma+1/k-\gamma}^{k/k-\gamma}) \\ &\quad d(\underline{x}_{k-\gamma}, \underline{s}_{k-\gamma}, \underline{\sigma}_{k-\gamma+1/k-\gamma}^{k/k-\gamma}) \end{aligned} \quad (X_E)$$

where for this instance, we introduce the notation

$$\underline{\sigma}_{k-\gamma+1/k-\gamma}^{k/k-\gamma} \stackrel{\text{Df}}{=} (\underline{\sigma}_{k/k-\gamma}, \dots, \underline{\sigma}_{k-\gamma+1, k-\gamma})$$

The density  $p(\underline{z}_k, \dots, \underline{z}_{k-\gamma+1}/\underline{x}_{k-\gamma})$  has the characteristic function

$$\varphi(\underline{\sigma}_{k-\gamma+1/k-\gamma}^{k/k-\gamma}) = \int \exp[i(\sum_{j=0}^{\gamma-1} \underline{\sigma}_{k-j/k-\gamma}^T \underline{z}_{k-j})] p(\underline{z}_{k-\gamma+1}^k/\underline{x}_{k-\gamma}) d(\underline{z}_{k-\gamma+1}^k) \quad (2.29)$$

where

$$\begin{aligned} p(\underline{z}_{k-\gamma+1}^k/\underline{x}_{k-\gamma}) &= \frac{1}{(2\pi)^{[(\gamma-1)m+n]}} \int [-i(\sum_{j=0}^{\gamma-2} \underline{\sigma}_{k-j/k-\gamma+1}^T \\ &\quad -i \underline{s}_v^T \underline{z}_{k-\gamma+1} - i \underline{s}_w^T \underline{x}_{k-\gamma+1})] \\ &\quad \varphi(\underline{\sigma}_{k-\gamma+2/k-\gamma+1}^{k/k-\gamma+1}) \varphi(\underline{s}_v) \varphi(\underline{s}_w) \\ &\quad d(\underline{\sigma}_{k-\gamma+2/k-\gamma+1}^{k/k-\gamma+1}, \underline{s}_v, \underline{x}_{k-\gamma+1}) \end{aligned} \quad (2.30)$$



The characteristic function for  $p(\underline{z}_k/\underline{x}_{k-1})$  is

$$\varphi(\underline{\sigma}_{k/k-1}) = \frac{1}{(2\pi)^{n+m}} \int \exp[-i(\underline{s}_v - \underline{\sigma}_{k/k-1})^T \underline{z}_k - i \underline{s}_w^T \underline{x}_k] \varphi(\underline{s}_v) \varphi(\underline{s}_w) d(\underline{z}_k, \underline{s}_v, \underline{s}_w) \quad (2.31)$$

The proof of this result is straightforward and shall be omitted.

PRECEDING PAGE BLANK NOT FILMED.

### CHAPTER THREE

#### THE LINEAR, TIME-DISCRETE STOCHASTIC CONTROL PROBLEM

The model of Chapter 1 shall be specialized to that of a linear system. The results presented in this chapter are not new but have been included to illustrate the application of the general theory of Chapter 2 to a problem of fundamental importance. It is believed that this discussion indicates the relative ease with which many of the most important results of the theory of linear systems are obtained using the Bayesian approach.

Assume that the plant is described by the linear, difference equation

$$\underline{x}_k = \Phi_{k,k-1} \underline{x}_{k-1} + \Gamma_{k,k-1} u_{k-1} + \underline{w}_{k-1} \quad (\text{I-L})$$

and the state is measured imperfectly according to

$$\underline{z}_k = H_k \underline{x}_k + \underline{v}_k \quad (\text{II-L})$$

The white noise sequences  $\{\underline{w}_j\}$  and  $\{\underline{v}_j\}$  shall be explicitly assumed to be gaussian as is the distribution of the initial state  $\underline{x}_0$ . The symbol L has been appended to the equation numbers to emphasize that the systems are linear.

The densities for plant noise  $\underline{w}_j$ , measurement noise  $\underline{v}_j$ , and initial condition  $\underline{x}_0$  are

$$p(\underline{w}_j) = [(2\pi)^n |Q_j|]^{-1/2} \exp - \frac{1}{2} \underline{w}_j^T Q_j^{-1} \underline{w}_j \quad (3.1)$$

$$p(\underline{v}_j) = [(2\pi)^m |R_j|]^{-1/2} \exp - \frac{1}{2} \underline{v}_j^T R_j^{-1} \underline{v}_j \quad (3.2)$$

$$p(\underline{x}_0) = [(2\pi)^n |M_0|]^{-1/2} \exp - \frac{1}{2} (\underline{x}_0 - \underline{a})^T M_0^{-1} (\underline{x}_0 - \underline{a}) \quad (3.3)$$

In order to write (3.1) - (3.3), it is necessary to assume that the covariance matrix of each distribution is positive-definite. If the matrix were singular, one could always consider the variable in the subspace spanned by the eigenvectors corresponding to the nonzero eigenvalues of the covariance matrix [10]. The covariance matrix of the transformed variable in the reduced space would be positive-definite. This difficulty can also be avoided by allowing the characteristic function to be the defining relation for the distribution and restricting consideration to this function [24]. The latter alternative shall be utilized in Section 3.1.

The control variables will be selected so that the expected value of the quadratic performance index

$$V_N = \sum_{i=1}^N (\underline{x}_i^T \underline{W}_i \underline{x}_i + \underline{u}_{i-1}^T \underline{W}_i \underline{u}_{i-1}) \quad (\text{III}_C\text{-L})$$

is minimized. The mean-square error criteria

$$E[(\hat{\underline{x}}_k - \underline{x}_k)^T (\hat{\underline{x}}_k - \underline{x}_k)] = \text{minimum} \quad (\text{III}_E\text{-L})$$

will be seen in Section 3.2 to be required in the solution of the control problem for the estimate of the state. Mean-square estimates are considered in Section 3.1 as a preliminary to the discussion of the control problem.

### 3.1 MINIMUM MEAN-SQUARE ESTIMATES

The model for the plant will be simplified in this section by the omission of the control terms. Then, the state evolves in accordance with

$$\underline{x}_k = \Phi_{k,k-1} \underline{x}_{k-1} + \underline{w}_{k-1} \quad (\text{I}_E\text{-L})$$

It was shown in Section 2.1 that the estimate resulting from the minimum mean-square error criteria is given by the conditional mean of the a posteriori density function. This is true for all three aspects (i.e., filtering, prediction, and smoothing) of the estimation problem and the solution to each shall be presented.

Two general results [10, 39] will be used in the discussion.

$$(1) \quad \frac{1}{(2\pi)^n} \int \exp[i \underline{s}^T \underline{x}] d\underline{x} = \delta(\underline{s}) \quad (3.4)$$

where  $\delta(\cdot)$  is the Dirac delta function

$$(2) \quad \int_{-\infty}^{\infty} \exp[\eta^T \underline{z} - \underline{z}^T A \underline{z}] d\underline{z} = \left( \frac{\pi^n}{|A|} \right)^{1/2} \exp\left[ \frac{1}{4} \eta^T A^{-1} \eta \right] \quad (3.5)$$

for any complex  $\eta$  and positive-definite  $A$ .

LEMMA 3.1: The a posteriori density  $p(\underline{x}_k / \underline{z}^k)$  for the system  $(I_E - L) - (II - L)$  is gaussian

$$p(\underline{x}_k / \underline{z}^k) = [(2\pi)^n |P_k|]^{-1/2} \exp - \frac{1}{2} \{ (\underline{x}_k - \hat{\underline{x}}_k)^T P_k^{-1} (\underline{x}_k - \hat{\underline{x}}_k) \} \quad (XI)$$

with mean value

$$\hat{\underline{x}}_k = \hat{\underline{x}}_k' + K_k (\underline{z}_k - H_k \hat{\underline{x}}_k') \quad (3.6)$$

where

$$\hat{\underline{x}}_k' = \hat{\Phi}_{k,k-1} \hat{\underline{x}}_{k-1} \quad (3.7)$$

$$K_k = P_k' H_k^T (H_k P_k' H_k^T + R_k)^{-1} \quad (3.8)$$

$$P_k' = \hat{\Phi}_{k,k-1} P_{k-1} \hat{\Phi}_{k,k-1}^T + Q_{k-1} \quad (3.9)$$

and covariance matrix

$$P_k = P'_k - K_k H_k P'_k \quad (3.10)$$

At  $t_0$  the mean value is

$$\hat{x}_0 = \underline{a} + K_0 (\underline{z}_0 - H_0 \underline{a}) \quad (3.11)$$

where

$$K_0 = M_0 H_0^T (H_0 M_0 H_0^T + R_0)^{-1} \quad (3.12)$$

and the covariance matrix is

$$P_0 = M_0 - K_0 H_0 M_0 \quad (3.13)$$

The equations described by this lemma constitute the so-called Kalman filter [25, 30]. Within the framework of the Bayesian approach, the proof has been found to be established most easily using the characteristic function formulation described in Section 2.3. Note again that with this approach, the covariance matrices need not be positive-definite.

Proof: Let us first establish the initial conditions (3.11) - (3.13). From (3.3) it follows that the characteristic function for  $\underline{x}_0$  is

$$\varphi(\underline{s}_m) = \exp\{i \underline{s}_m^T \underline{a} - \frac{1}{2} \underline{s}_m^T M_0 \underline{s}_m\} \quad (3.14)$$

and from (II-L) and (3.2) the characteristic function of  $\underline{z}_0$  given  $\underline{x}_0$  is

$$\varphi(\underline{s}_v) = \exp\{i \underline{s}_v^T H_0 \underline{x}_0 - \frac{1}{2} \underline{s}_v^T R_0 \underline{s}_v\} \quad (3.15)$$

Substitute (3.14) and (3.15) into (2.25) and let

$$k_0 \stackrel{\text{Df}}{=} \frac{1}{(2\pi)^{n+m} p(\underline{z}_0)}$$

$$\varphi(\underline{s}_0) = k_0 \int \exp[-i(\underline{s}_m - \underline{s}_0 - H_{0-v}^T \underline{s}_0) \underline{x}_0 - i \underline{s}_v^T \underline{z}_0 + i \underline{s}_m^T \underline{a} - \frac{1}{2} \underline{s}_m^T M_{0-m} \underline{s}_m - \frac{1}{2} \underline{s}_v^T R_{0-v} \underline{s}_v] d(\underline{s}_m, \underline{s}_v, \underline{x}_0)$$

Integrate with respect to  $\underline{x}_0$  and use (3.4). Then

$$\varphi(\underline{s}_0) = k_0 \int \delta(H_{0-v}^T \underline{s}_v + \underline{s}_0 - \underline{s}_m) \exp[-i \underline{s}_v^T \underline{z}_0 + i \underline{s}_m^T \underline{a} - \frac{1}{2} \underline{s}_m^T M_{0-m} \underline{s}_m - \frac{1}{2} \underline{s}_v^T R_{0-v} \underline{s}_v] d(\underline{s}_m, \underline{s}_v)$$

After integrating with respect to  $\underline{s}_m$ , this becomes

$$\varphi(\underline{s}_0) = k_0 \exp[i \underline{s}_0^T \underline{a} - \frac{1}{2} \underline{s}_0^T M_{0-0} \underline{s}_0] \int \exp\{\underline{s}_v^T [-i(\underline{z}_0 - H_{0-a}) - H_{0-v} M_{0-0} \underline{s}_0] - \frac{1}{2} \underline{s}_v^T (H_{0-v} M_{0-v} H_{0-v}^T + R_{0-v}) \underline{s}_v\} d\underline{s}_v$$

Using (3.5) and evaluating  $p(\underline{z}_0)$ , this reduces to

$$\varphi(\underline{s}_0) = \exp\{i \underline{s}_0^T [\underline{a} + K_{0-v} (\underline{z}_0 - H_{0-a})] - \frac{1}{2} \underline{s}_0^T [M_{0-v} - K_{0-v} H_{0-v} M_{0-0}] \underline{s}_0\} \quad (3.16)$$

But (3.16) is the characteristic function equivalent of (XI) with mean and covariance described by (3.11) - (3.13).

To verify (3.6) - (3.10) assume that the lemma is true for  $t_{k-1}$  and form

$\varphi(\underline{s}_{k/k-1})$ . From (I<sub>E</sub>-L) and (3.11)

$$\varphi(\underline{s}_w) = \exp\{i \underline{s}_w^T \hat{\underline{x}}_{k,k-1} - \frac{1}{2} \underline{s}_w^T Q_{k-1} \underline{s}_w\} \quad (3.17)$$

Substitute (3.17) and  $\varphi(\underline{s}_{k-1})$  into (2.23). It follows in a straightforward manner that

$$\varphi(\underline{s}_{k/k-1}) = \exp[i \underline{s}_{k/k-1}^T \hat{\underline{x}}_k - \frac{1}{2} \underline{s}_{k/k-1}^T P'_k \underline{s}_{k/k-1}] \quad (3.18)$$

where  $\hat{x}'_k$  and  $P'_k$  are defined by (3.7) and (3.9). Note that this provides a solution of the one-stage prediction problem.

The proof of (XI) with (3.6), (3.8), and (3.10) proceeds in a manner that is identical with that used to derive (3.11) - (3.13) except that  $\hat{x}'_k$  and  $P'_k$  replace  $\underline{a}$  and  $M_o$ .

Q. E. D.

It can be proven immediately from  $(IX_E)$  that the prediction problem has the following solution.

LEMMA 3.2: The a posteriori density  $p(\underline{x}_k/\underline{z}^{k-\gamma})$ ,  $\gamma > 0$ , for the system  $(I_E-L)$  -  $(II_E-L)$  is gaussian

$$p(\underline{x}_k/\underline{z}^{k-\gamma}) = [(2\pi)^n |P_{k/k-\gamma}|]^{-1/2} \exp -\frac{1}{2} \{ (\underline{x}_{k/k-\gamma} - \hat{\underline{x}}_{k/k-\gamma})^T P_{k/k-\gamma}^{-1} (\underline{x}_{k/k-\gamma} - \hat{\underline{x}}_{k/k-\gamma}) \} \quad (XII)$$

with mean value

$$\hat{\underline{x}}_{k/k-\gamma} = \Phi_{k,k-\gamma} \hat{\underline{x}}_{k-\gamma} \quad (3.19)$$

and covariance computed recursively from

$$P_{k/k-\gamma} = \Phi_{k,k-1} P_{k-1/k-\gamma} \Phi_{k,k-1}^T + Q_{k-1} \quad (3.20)$$

where

$$P_{k-\gamma+1/k-\gamma} = \Phi_{k-\gamma+1,k-\gamma} P_{k-\gamma} \Phi_{k-\gamma+1,k-\gamma}^T + Q_{k-\gamma}$$

The proof was established to a major extent in the derivation of (3.18).

The remainder of the proof shall be omitted.

The solution of the smoothing problem requires more involved algebraic manipulations than were required for the prediction and smoothing problems.

The equations stated in the following lemma were first derived by Rauch [36].

LEMMA 3.3: The a posteriori density  $p(\underline{x}_{k-\gamma}/\underline{z}^k)$ ,  $\gamma > 0$ , for the system  $(I_E - L) - (II_E - L)$  is gaussian

$$p(\underline{x}_{k-\gamma}/\underline{z}^k) = [(2\pi)^n |P_{k-\gamma/k}|]^{-1/2} \exp - \frac{1}{2} \{ (\underline{x}_{k-\gamma/k} - \hat{\underline{x}}_{k-\gamma/k})^T P_{k-\gamma/k}^{-1} (\underline{x}_{k-\gamma/k} - \hat{\underline{x}}_{k-\gamma/k}) \} \quad (XIII)$$

with mean value

$$\hat{\underline{x}}_{k-\gamma/k} = \hat{\underline{x}}_{k-\gamma} + C_{k-\gamma} [\hat{\underline{x}}_{k-\gamma+1/k} - \Phi_{k-\gamma+1, k-\gamma} \hat{\underline{x}}_{k-\gamma}] \quad (3.21)$$

where

$$C_{k-\gamma} = P_{k-\gamma} \Phi_{k-\gamma+1/k-\gamma}^T P_{k-\gamma+1}^{-1} \quad (3.22)$$

and covariance

$$P_{k-\gamma/k} = P_{k-\gamma} + C_{k-\gamma} (P_{k-\gamma+1/k} - P_{k-\gamma+1}') C_{k-\gamma}^T \quad (3.23)$$

Proof: These relations shall only be verified for a one and two-stage processes.

Consider a one-stage problem (i.e.,  $\gamma = 1$ ). Then, substitute (3.15) and (3.17) into (2.31). This yields

$$\begin{aligned} \varphi(\underline{\sigma}_{k/k-1}) &= \frac{1}{(2\pi)^{n+m}} \int \exp[-i(\underline{s}_v - \underline{\sigma}_{k/k-1})^T \underline{z}_k - i(\underline{s}_w - H_k^T \underline{s}_v)^T \underline{x}_k] \\ &\quad - \frac{1}{2} (\underline{s}_v^T R_{k-v} \underline{s}_v + \underline{s}_w^T Q_{k-1} \underline{s}_w) d(\underline{z}_k, \underline{x}_k, \underline{s}_v, \underline{s}_w) \end{aligned} \quad (3.24)$$



Integrate with respect to  $\underline{z}_k$  and  $\underline{x}_k$ . These integrations will introduce the delta functions  $\delta(\underline{s}_v - \underline{\sigma}_{k/k-1})$  and  $\delta(\underline{s}_w - H_k^T \underline{s}_v)$ . Next, integrate relative to  $\underline{s}_w$  and then with respect to  $\underline{s}_v$ . This leads to

$$\varphi(\underline{\sigma}_{k/k-1}) = \frac{1}{(2\pi)^{n+m}} \exp[i \underline{\sigma}_{k/k-1}^T H_k^T \hat{\Phi}_{k,k-1} \underline{x}_{k-1} - \frac{1}{2} \underline{\sigma}_{k/k-1}^T (H_k Q_{k-1} H_k^T + R_k) \underline{\sigma}_{k/k-1}] \quad (3.25)$$

The characteristic function for  $\underline{x}_{k-1}$  given  $\underline{z}^k$  according to  $(X_E)$  is

$$\varphi(\underline{s}_{k-1/k}) = k_{k-1/k} \int \exp[-i(\underline{s}_{k-1} - \underline{s}_{k-1/k})^T \underline{x}_{k-1} - i \underline{\sigma}_{k/k-1}^T \underline{z}_k] \varphi(\underline{s}_{k-1}) \varphi(\underline{\sigma}_{k/k-1}) d(\underline{x}_{k-1}, \underline{s}_{k-1}, \underline{\sigma}_{k/k-1})$$

where

$$k_{k-1/k} \stackrel{\text{Df}}{=} \frac{1}{2\pi^{[(\gamma+1)m+n]} p(\underline{z}_k / \underline{z}^{k-1})}$$

From (3.25) and (XI), this becomes

$$\begin{aligned} \varphi(\underline{s}_{k-1/k}) = k_{k-1/k} \int \exp[ & i(\underline{s}_{k-1/k} - \underline{s}_{k-1} + \hat{\Phi}_{k,k-1}^T H_k^T \underline{\sigma}_{k/k-1})^T \underline{x}_{k-1} \\ & - i \underline{\sigma}_{k/k-1}^T \underline{z}_k + i \underline{s}_{k-1}^T \hat{\underline{x}}_{k-1} - \frac{1}{2} \underline{s}_{k-1}^T P_{k-1} \underline{s}_{k-1} \\ & - \frac{1}{2} \underline{\sigma}_{k/k-1}^T (H_k Q_{k-1} H_k^T + R_k) \underline{\sigma}_{k/k-1}] d(\underline{x}_{k-1}, \underline{\sigma}_{k/k-1}, \underline{s}_{k-1}) \end{aligned}$$

Integration with respect to  $\underline{x}_{k-1}$  introduces the delta function  $\delta(\underline{s}_{k-1/k} - \underline{s}_{k-1} + \hat{\Phi}_{k,k-1}^T H_k^T \underline{\sigma}_{k/k-1})$ . This is removed by integrating with respect to  $\underline{s}_{k-1}$ .

Then

$$\begin{aligned}
\varphi(\underline{s}_{k-1/k}) &= k_{k-1/k} \exp[i \underline{s}_{k-1/k}^T \hat{\underline{x}}_{k-1} - \frac{1}{2} \underline{s}_{k-1/k}^T P_{k-1} \underline{s}_{k-1/k}] \\
&\int \exp\{[i(H_k \hat{\underline{\phi}}_{k,k-1} \hat{\underline{x}}_{k-1} - \underline{z}_k)^T - \underline{s}_{k-1/k}^T P_{k-1} \hat{\underline{\phi}}_{k,k-1}^T H_k^T] \underline{\sigma}_{k/k-1} \\
&- \frac{1}{2} \underline{\sigma}_{k/k-1}^T (H_k P_k' H_k^T + R_k) \underline{\sigma}_{k/k-1}\} d\underline{\sigma}_{k/k-1} \quad (3.26)
\end{aligned}$$

This integral is evaluated by inspection by applying (3.5). After the constant  $k_{k-1/k}$  is determined, (3.26) becomes

$$\varphi(\underline{s}_{k-1/k}) = \exp[i \underline{s}_{k-1/k}^T \hat{\underline{x}}_{k-1/k} - \frac{1}{2} \underline{s}_{k-1/k}^T P_{k-1/k} \underline{s}_{k-1/k}] \quad (3.27)$$

The mean value is

$$\hat{\underline{x}}_{k-1/k} = \hat{\underline{x}}_{k-1} + K_{k-1/k} (\underline{z}_k - H_k \hat{\underline{\phi}}_{k,k-1} \hat{\underline{x}}_{k-1}) \quad (3.28)$$

where

$$K_{k-1/k} = P_{k-1} \hat{\underline{\phi}}_{k,k-1}^T H_k^T (H_k P_k' H_k^T + R_k)^{-1} \quad (3.29)$$

and the covariance is

$$P_{k-1/k} = P_{k-1} - K_{k-1/k} H_k \hat{\underline{\phi}}_{k,k-1} P_{k-1} \quad (3.30)$$

Equations (3.28) - (3.30) do not appear to have the form described by the lemma, but it will be shown that they are equivalent.

To prove the equivalence of (3.28) - (3.30) and (3.21) - (3.23) observe that (3.30) can be written as

$$P_{k-1/k} = P_{k-1} - P_{k-1} \hat{\underline{\phi}}_{k,k-1}^T P_k'^{-1} K_k H_k \hat{\underline{\phi}}_{k,k-1} P_{k-1}$$

But from (3.10)

$$K_k H_k = I - P_k P_k'^{-1}$$

so

$$P_{k-1/k} = P_{k-1} + P_{k-1} \Phi_{k,k-1}^T P_k'^{-1} (P_k - P_k') P_k'^{-1} \Phi_{k,k-1} P_{k-1}$$

But this is in accord with (3.23) when  $\gamma = 1$  and the definition of  $C_{k-1}$  is introduced. (3.28) reduces to (3.21) by recalling from (3.6) - (3.8) that

$$K_k (z_k - H_k \Phi_{k,k-1} \hat{x}_{k-1}) = \hat{x}_k - \Phi_{k,k-1} \hat{x}_{k-1}$$

This allows (3.28) to be written as

$$\hat{x}_{k-1/k} = \hat{x}_{k-1} + P_{k-1} \Phi_{k,k-1}^T P_k'^{-1} (\hat{x}_k - \Phi_{k,k-1} \hat{x}_{k-1})$$

which completes the proof for  $\gamma = 1$ .

The derivation of the smoothing density for  $\gamma > 1$  becomes considerably more involved. For  $\gamma = 2$ , one finds that  $\varphi(\sigma_{k/k-2}, \sigma_{k-1/k-2})$  is

$$\begin{aligned} \varphi(\sigma_{k/k-2}, \sigma_{k-1/k-2}) &= \frac{1}{(2\pi)^{m+n}} \\ &\exp[i(\Phi_{k,k-1}^T H_{k-2}^T \sigma_{k/k-1} + H_{k-1}^T \sigma_{k-1/k-2})^T \\ &\quad \Phi_{k-1,k-2} x_{k-2}] \\ &\exp\left\{-\frac{1}{2} [\sigma_{k/k-2}^T \quad \sigma_{k-1/k-2}^T] \right. \\ &\quad \left. \begin{bmatrix} \bar{N}_{k/k-2} & A_{k-2} \\ A_{k-2}^T & N_{k-1/k-2} \end{bmatrix} \begin{bmatrix} \sigma_{k/k-2} \\ \sigma_{k-1/k-2} \end{bmatrix} \right\} \quad (3.31) \end{aligned}$$

where

$$N_{k/k-2} \stackrel{\text{Df}}{=} H_k (\Phi_{k,k-1} Q_{k-2} \Phi_{k,k-1}^T + Q_{k-1}) H_k^T + R_k$$

$$N_{k-1/k-2} \stackrel{\text{Df}}{=} H_{k-1} Q_{k-2} H_{k-2}^T + R_{k-1}$$

$$A_{k-2} \stackrel{Df}{=} H_k \Phi_{k,k-1} Q_{k-2} H_{k-1}^T$$

Then, after considerable manipulation, one obtains

$$\varphi(\underline{s}_{k-2/k}) = \exp[i \underline{s}_{k-2/k}^T \hat{\underline{x}}_{k-2/k} - \frac{1}{2} \underline{s}_{k-2/k}^T P_{k-2/k} \underline{s}_{k-2/k}] \quad (3.32)$$

where

$$\begin{aligned} \hat{\underline{x}}_{k-2/k} &= \hat{\underline{x}}_{k-2} + P_{k-2} \Phi_{k-1,k-2}^T P_{k-1}'^{-1} [P_{k-1} \Phi_{k,k-1}^T H_k^T (H_k P_k' H_k^T + R_k)^{-1} \\ &\quad (\underline{z}_k - H_k \Phi_{k,k-2} \hat{\underline{x}}_{k-2}) + P_{k-1/k} P_{k-1}^{-1} P_{k-1}' H_{k-1}^T \\ &\quad (H_{k-1} P_{k-1}' H_{k-1}^T + R_{k-1})^{-1} (\underline{z}_k - H_{k-1} \Phi_{k-1,k-2} \hat{\underline{x}}_{k-2})] \end{aligned} \quad (3.33)$$

and

$$\begin{aligned} P_{k-2/k} &= P_{k-2} - P_{k-2} \Phi_{k-1,k-2}^T [P_{k-1}'^{-1} P_{k-1/k} P_{k-1}^{-1} P_{k-1}' H_{k-1}^T \\ &\quad (H_{k-1} P_{k-1}' H_{k-1}^T + R_{k-1})^{-1} H_{k-1} \\ &\quad + P_{k-1}' P_{k-1} \Phi_{k,k-1}^T H_k^T (H_k P_k' H_k^T + R_k)^{-1} H_k \Phi_{k,k-1}] \\ &\quad \Phi_{k-1,k-2} P_{k-2} \end{aligned} \quad (3.34)$$

(3.33) and (3.34) can be shown to be equivalent to (3.21) and (3.23).

Q. E. D.

The preceding lemmas provide the complete solution of the estimation problem for linear systems. The filter equations will be required in the discussion of the stochastic control problem as presented in the next section.

### 3.2 THE LINEAR FEEDBACK CONTROL LAW

In this section the control law for the system (I-L) - (II-L) is derived under the constraint that the control minimizes the expected value of the performance index (III<sub>C</sub>-L). Before dealing with this problem, a result from the theory of optimal control of deterministic systems shall be stated.

Suppose that the plant is described by

$$\underline{x}_k = \Phi_{k,k-1} \underline{x}_{k-1} + \Gamma_{k,k-1} u_{k-1} \quad (3.35)$$

and that  $\underline{x}_k$  is known at each sampling time  $t_k$ . The control policy that minimizes (III<sub>C</sub>-L) under these constraints is given by the following lemma [40].

LEMMA 3.4: The optimal control  $u_k^0$  for the system (3.35) and performance index (III<sub>C</sub>-L) is described by

$$u_{N-k-1}^0 = -\Lambda_{N-k} \Phi_{N-k,N-k-1} \underline{x}_{N-k-1} \quad (XIV)$$

where

$$\Lambda_{N-k} = (\Gamma_{N-k,N-k-1}^T \Pi'_{N-k} \Gamma_{N-k,N-k-1} + W_{N-k-1}^U)^{-1} \Gamma_{N-k,N-k-1}^T \Pi'_{N-k} \quad (3.36)$$

$$\Pi'_{N-k} = \Phi_{N-k+1,N-k}^T \Pi_{N-k+1} \Phi_{N-k+1,N-k} + W_{N-k}^X \quad (3.37)$$

$$\Pi_{N-k} = \Pi'_{N-k} - \Pi'_{N-k} \Gamma_{N-k,N-k-1} \Lambda_{N-k} \quad (3.38)$$

For  $k = 0$ , the  $\Pi_{N+1}$  appearing in (3.37) is taken to be identically zero.

It is interesting to observe the similarity of (3.36) - (3.38) to the gain and covariance matrices (3.8) - (3.10) of the optimal filter. This similarity has

been recognized by Kalman and formalized in terms of a "Duality Principle" [26, 40].

The control law (XIV) has been included because it plays a fundamental role in the solution of the stochastic control law. This problem has the solution described in the following statement.

**SEPARATION PRINCIPLE:** For the model described by (I-L), (II-L), and (III<sub>C</sub>-L), the optimal stochastic control law is described by

$$\underline{u}_{N-k-1}^0 = -\Lambda_{N-k}^{\Phi} \hat{\underline{x}}_{N-k-1}$$

where  $\Lambda_{N-k}$  is defined by (3.36) - (3.38). The  $\hat{\underline{x}}_{N-k-1}$  is the minimum mean-square estimate of the state  $\underline{x}_{N-k-1}$  as obtained from the measurement data  $\underline{z}^{N-k-1}$ . In obtaining the estimate, the  $\underline{u}^{N-k-2}$  is treated as a deterministic function.

**Proof:** The proof of this principle is obtained through the direct application of the lemma of Section 2.1.2. Consider the last stage.

$$\begin{aligned} \bar{\mathcal{J}}_N &= \int (\underline{x}_N^T \underline{W}_N^X \underline{x}_N + \underline{u}_{N-1}^T \underline{W}_{N-1}^U \underline{u}_{N-1}) \\ &\quad \delta(\underline{x}_N - \Phi_{N,N-1} \underline{x}_{N-1} - \Gamma_{N,N-1} \underline{u}_{N-1} - \underline{w}_{N-1}) \\ &\quad p(\underline{w}_{N-1}) d(\underline{x}_N, \underline{w}_{N-1}) \end{aligned} \quad (3.39)$$

Carry out the indicated integrations. This yields

$$\begin{aligned} \bar{\mathcal{J}}_N &= \underline{x}_{N-1}^T \Phi_{N,N-1}^T \underline{W}_N^X \Phi_{N,N-1} \underline{x}_{N-1} + 2 \underline{x}_{N-1}^T \Phi_{N,N-1}^T \underline{W}_N^X \Gamma_{N,N-1} \underline{u}_{N-1} \\ &\quad + \underline{u}_{N-1}^T (\underline{W}_{N-1}^U + \Gamma_{N,N-1}^T \underline{W}_N^X \Gamma_{N,N-1}) \underline{u}_{N-1} + \text{trace} [\underline{W}_N^Q \underline{Q}_{N-1}] \end{aligned} \quad (3.40)$$

The control  $\underline{u}_{N-1}$  is to be chosen to minimize the conditional expectation of  $\bar{\mathcal{J}}_N$ .

$$\begin{aligned}
E[\bar{z}_N / \underline{z}^{N-1}] &= E[\underline{x}_{N-1}^T \Phi_{N,N-1}^T W_{N,N-1}^X \underline{x}_{N-1}] \\
&+ 2 \underline{u}_{N-1}^T \Gamma_{N,N-1}^T W_{N,N-1}^X \hat{\underline{x}}_{N-1} \\
&+ \underline{u}_{N-1}^T (W_{N-1}^U + \Gamma_{N,N-1}^T W_{N,N-1}^X) \underline{u}_{N-1} \\
&+ \text{trace} [W_N^X Q_{N-1}]
\end{aligned} \tag{3.41}$$

where we used the fact that

$$\hat{\underline{x}}_{N-1} = E[\underline{x}_{N-1} / \underline{z}^{N-1}]$$

Since  $\hat{\underline{x}}_{N-1}$  assumes that  $\underline{z}^{N-1}$  is given, the controls  $\underline{u}^{N-2}$  are known and can be treated as deterministic forcing functions. Then, from Reference 30 we know that the error in the estimate is independent of a known function.

It follows immediately from (3.41) that the control that minimizes  $E[\bar{z}_N / \underline{z}^{N-1}]$  is

$$\underline{u}_{N-1}^0 = -(\Gamma_{N,N-1}^T W_{N,N-1}^X \Gamma_{N,N-1} + W_{N-1}^U)^{-1} \Gamma_{N,N-1}^T W_{N,N-1}^X \underline{x}_{N-1} \tag{3.42}$$

Let

$$\Pi_N' \stackrel{\text{Df}}{=} W_N^X$$

and

$$\Lambda_N \stackrel{\text{Df}}{=} (\Gamma_{N,N-1}^T \Pi_N' \Gamma_{N,N-1} + W_{N-1}^U)^{-1} \Gamma_{N,N-1}^T \Pi_N'$$

Then, (3.42) satisfies the statement of the Separation Principle for the last stage.

Consider a two-stage problem. The control for the last stage is given by (3.42) and, using it, one can form  $\bar{z}_N^0$ .

$$\begin{aligned}
\bar{J}_N^0 &= \underline{x}_{N-1}^T \Phi_{N,N-1}^T \Pi_N' \Phi_{N,N-1} \underline{x}_{N-1} \\
&\quad - 2 \underline{x}_{N-1}^T \Phi_{N,N-1}^T \Pi_N' \Gamma_{N,N-1} \Lambda_N \Phi_{N,N-1} \hat{\underline{x}}_{N-1} \\
&\quad + \hat{\underline{x}}_{N-1}^T \Phi_{N,N-1}^T \Pi_N' \Gamma_{N,N-1} \Lambda_N \Phi_{N,N-1} \hat{\underline{x}}_{N-1} \\
&\quad + \text{trace} [W_N^X Q_{N-1}]
\end{aligned} \tag{3.43}$$

At this point recall that the estimate can be stated as

$$\hat{\underline{x}}_{k-1} = \tilde{\underline{x}}_{k-1} + \underline{x}_{k-1}$$

Use this relation to eliminate  $\hat{\underline{x}}_{N-1}$  in (3.43). After regrouping terms, the  $\bar{J}_N^0$  is seen to be

$$\begin{aligned}
\bar{J}_N^0 &= \underline{x}_{N-1}^T \Phi_{N,N-1}^T \Pi_N' \Phi_{N,N-1} \underline{x}_{N-1} \\
&\quad + \tilde{\underline{x}}_{N-1}^T \Phi_{N,N-1}^T \Pi_N' \Gamma_{N,N-1} \Lambda_N \Phi_{N,N-1} \tilde{\underline{x}}_{N-1} \\
&\quad + \text{trace} [W_N^X Q_{N-1}]
\end{aligned} \tag{3.44}$$

where

$$\Pi_N = \Pi_N' - \Pi_N' \Gamma_{N,N-1} \Lambda_N$$

The  $\Pi_N$  agrees with (3.38). The cost associated with the optimal control is

$$\begin{aligned}
E[V_1^0] &= E\{E[\bar{J}_N^0 / \underline{z}^{N-1}]\} \\
&= E\{E[\underline{x}_{N-1}^T \Phi_{N,N-1}^T \Pi_N' \Phi_{N,N-1} \underline{x}_{N-1} | \underline{z}^{N-1}] \\
&\quad + \text{trace} \Phi_{N,N-1}^T \Pi_N' \Gamma_{N,N-1} \Lambda_N \Phi_{N,N-1} E[\tilde{\underline{x}}_{N-1} \tilde{\underline{x}}_{N-1}^T | \underline{z}^{N-1}]\} \\
&\quad + \text{trace} [W_N^X Q_{N-1}]
\end{aligned}$$



At this juncture it is important to recognize that the conditional covariance  $E[\tilde{x}_{N-1}\tilde{x}_{N-1}^T | z^{N-1}]$  is independent of the control vectors  $u^{N-1}$  and the measurements  $z^{N-1}$ . This follows from the results in Section 3.1. Because of this fact, only the first term must be considered in determining the optimal control for earlier times. (This aspect is discussed further in Section 4.3.) Since the term  $\text{trace} \{ \Phi_{N,N-1}^T \Pi'_{N,N-1} \Gamma_{N,N-1} \Lambda_{N,N-1} \Phi_{N,N-1} P_{N-1} + W_{N,N-1}^X Q_{N,N-1} \}$  has no bearing on the selection of the control policy, it will be neglected and the  $\bar{Z}_N^0$  will be redefined as

$$\bar{Z}_N^0 = x_{N-1}^T \Phi_{N,N-1}^T \Pi_{N,N-1} \Phi_{N,N-1} x_{N-1} \quad (3.45)$$

and it follows immediately from  $(V_C)$  that

$$\bar{Z}'_{N-1} = \bar{Z}_N^0$$

Thus,

$$\begin{aligned} \bar{Z}_{N-1} &= \int (x_{N-1}^T \Pi'_{N-1} x_{N-1} + u_{N-2}^T W_{N-2}^U u_{N-2}) \\ &\quad \delta(x_{N-1} - \Phi_{N-1,N-2} x_{N-2} - \Gamma_{N-1,N-2} u_{N-2} - w_{N-2}) \\ &\quad p(w_{N-2}) d(x_{N-1}, w_{N-2}) \end{aligned} \quad (3.46)$$

where

$$\Pi'_{N-1} \stackrel{\text{Df}}{=} W_{N-1}^X + \Phi_{N,N-1}^T \Pi_{N,N-1} \Phi_{N,N-1}$$

It follows without difficulty that the control  $u_{N-2}$  that minimizes  $E[\bar{Z}_{N-1}/z^{N-2}]$  is

$$u_{N-2}^0 = -\Lambda_{N-1} \Phi_{N-1,N-2} \hat{x}_{N-2}$$

where  $\Lambda_{N-1}$  is defined by (3.36) - (3.38). The proof for any  $k$  assuming that the  $\bar{Z}_{N-k-2}^0$  is (again, retaining only those terms that depend upon the control)

$$\bar{Z}_{N-k+2}^0 = \bar{x}_{N-k+1}^T \bar{\Phi}_{N-k+2, N-k+1}^T \Pi_{N-k+2} \bar{\Phi}_{N-k+2, N-k+1} \bar{x}_{N-k+1}$$

is obtained directly from  $(V_C)$ .

Q. E. D.

This completes the solution of the optimal stochastic control problem for the linear system (I-L) - (II-L) and the quadratic performance index  $(III_C-L)$ . By necessity, the discussion has been restricted to the most important aspects of the problem. The reader is directed to References 40 to 44 for a more detailed examination of the linear problem.

## CHAPTER FOUR

### A GENERALIZATION OF THE KALMAN FILTER

In this chapter, the perturbative Bayesian scheme described in Section 1.3 is applied to the problem of determining an approximation to the a posteriori density function associated with a nonlinear system. In compliance with the aforementioned technique, the form of the density must be specified. It shall be required to be gaussian for all  $k$ . This leads to a natural generalization of the Kalman filter and suggests several interesting conclusions.

The approximation that is described in this chapter represents a generalization of a result obtained by Aoki [33]. Results obtained by other investigators also indicate that the Kalman filter does not represent the most general gaussian approximation. This problem has been considered for time-continuous systems by Bucy [20], Bass et al [21], and Fisher [23]. Jazwinski [45] has dealt with cases that involve discrete measurement data. His result has the disadvantage that it does not reduce to the Kalman filter when the nonlinear effects are set equal to zero. It is shown in Section 4.2 that the equations derived here do reduce to Kalman's relations.

The general result is stated in Section 4.1, and an outline of the derivation is presented. Several interesting conclusions follow from this result, and these aspects are discussed in Section 4.2. The control of a system described by a linear plant and nonlinear measurements is discussed in the light of this approximation, and it is suggested that the Separation Principle is no longer valid for this system.

The filter resulting from this approximation is utilized in Chapters 5 and 7 to determine its behavior relative to linear and other nonlinear filters. The results that are obtained, particularly those in Chapter 7, suggest that one must approach the problem of approximating the a posteriori density with caution because it appears that the estimates provided by this filter are biased. This undesirable feature is discussed in more detail below. Another gaussian approximation is discussed in Chapters 6 and 7.

#### 4.1 AN A POSTERIORI GAUSSIAN DENSITY FOR FILTERING OF NONLINEAR SYSTEMS

Consider a system in which the state  $\underline{x}_k$  evolves according to the non-linear difference equation

$$\underline{x}_k = \underline{f}_k(\underline{x}_{k-1}) + \underline{w}_{k-1} \quad (\text{I-N})$$

where  $\underline{x}_k$  is n-dimensional. The additive noise  $\underline{w}_{k-1}$  is a gaussian sequence with mean and covariance

$$E[\underline{w}_j] = \underline{0} \quad \text{for all } j$$

$$E[\underline{w}_j \underline{w}_k^T] = Q_k \delta_{kj}$$

Note that no control terms are included in (I-N)

The measurement data  $\underline{z}_k$  are described by

$$\underline{z}_k = \underline{h}_k(\underline{x}_k) + \underline{v}_k \quad (\text{II-N})$$

where  $\underline{z}_k$  is m-dimensional. The additive noise  $\underline{v}_k$  is a gaussian sequence with mean and covariance

$$E[\underline{v}_j] = \underline{0} \quad \text{for all } j$$

$$E[\underline{v}_j \underline{v}_k^T] = R_k \delta_{kj}$$

The sequences  $\{\underline{v}_k\}$  and  $\{\underline{w}_k\}$  are assumed to be independent. That is

$$E[\underline{v}_k \underline{w}_j^T] \equiv 0 \quad \text{for all } k, j.$$

The initial state  $\underline{x}_0$  is taken to be a gaussian random variable with mean and covariance

$$E[\underline{x}_0] = \underline{a}$$

$$E[\underline{x}_0 \underline{x}_0^T] = M_0$$

Also, the  $\underline{x}_0$  is independent of the noise sequences.

$$E[\underline{x}_0 \underline{v}_j^T] \equiv 0 \equiv E[\underline{x}_0 \underline{w}_k^T]$$

The covariance matrices  $\{R_k\}$ ,  $\{Q_k\}$ , and  $M_0$  shall be assumed to be positive-definite in much of the succeeding presentation, but this is not a severe restriction. If any of these matrices were singular, an appropriate linear transformation would yield random variables of smaller dimension that have positive-definite covariance matrices and the derivation would be carried out in terms of the new variables. Further, the restriction can be seen to be relaxed in the final relations that are obtained for the estimation policy.

The noise has been assumed to be additive in (I-N) and (II-N) in order to simplify the densities  $p(\underline{x}_k / \underline{x}_{k-1})$  and  $p(\underline{z}_k / \underline{x}_k)$ . If non-additive noise were assumed, it would be necessary to introduce the Jacobians of  $\underline{f}_k$  and  $\underline{h}_k$  with the concomitant complications.

The procedure described in Section 1.3 shall be used to approximate the a posteriori density function  $p(\underline{x}_k/\underline{z}^k)$ . It will be assumed that nominal values  $\underline{x}_{k-1}^*$  are available that will permit the conditional density  $p(\underline{x}_k/\underline{z}^k)$  to be gaussian for all  $k$ . Taylor series expansions of the  $\underline{f}_k$  and  $\underline{h}_k$  will be introduced using the nominal values of the state. This procedure leads to a generalization of the results for linear systems (i.e., of the Kalman filter).

Before proceeding further, let us introduce some of the notations that will appear. Let

$$F_k \stackrel{\text{Df}}{=} \left( \frac{\partial \underline{f}_{k+1}}{\partial \underline{x}_k} \right)^* = \begin{pmatrix} \frac{\partial f_{k+1}^1}{\partial x_k^1} & \cdots & \frac{\partial f_{k+1}^1}{\partial x_k^n} \\ \vdots & & \vdots \\ \frac{\partial f_{k+1}^n}{\partial x_k^1} & & \frac{\partial f_{k+1}^n}{\partial x_k^n} \end{pmatrix} \quad \text{evaluated with } \underline{x}_k^*$$

with the superscripts denote the component of the vector. The first partial derivatives of  $\underline{h}_{k+1}$  with respect to  $\underline{x}_{k+1}$  are

$$H_{k+1} \stackrel{\text{Df}}{=} \left( \frac{\partial \underline{h}_{k+1}}{\partial \underline{x}_{k+1}} \right)^*$$

The derivatives are evaluated with

$$\underline{\xi}_{k+1} \stackrel{\text{Df}}{=} \underline{f}_{k+1}(\underline{x}_k^*)$$

Second partial derivatives are also used. The second partials of the  $i^{\text{th}}$  component of  $\underline{f}_{k+1}$  and  $\underline{h}_{k+1}$  are denoted as

$$G_k^i \stackrel{\text{Df}}{=} \left( \frac{\partial^2 f_{k+1}^i}{\partial \underline{x}_k \partial \underline{x}_k} \right)^*$$

and

$$J_{k+1}^i \stackrel{\text{Df}}{=} \left( \frac{\partial^2 h_{k+1}^i}{\partial \underline{x}_{k+1} \partial \underline{x}_{k+1}} \right)^*$$

The  $G_k^i$  and  $J_{k+1}^i$  are evaluated with  $\underline{x}_k^*$  and  $\xi_{k+1}$ , respectively.

Finally, the perturbations in the state and measurement vectors are

$$\delta \underline{x}_k \stackrel{\text{Df}}{=} \underline{x}_k - \underline{x}_k^*$$

and

$$\delta \underline{z}_{k+1} \stackrel{\text{Df}}{=} \underline{z}_{k+1} - \underline{h}_{k+1}(\xi_{k+1})$$

With this introduction, we make the following assertion.

LEMMA 4.1: Suppose that the  $f_k$  and  $h_k$  of (I-N) and (II-N) have at least continuous second partial derivatives. Then, assuming that there exists some nominal value  $\underline{x}_k^*$  of the state that is a sufficiently good approximation, the a posteriori density  $p(\underline{x}_{k+1}/\underline{z}^{k+1})$  can be written

$$p(\underline{x}_{k+1}/\underline{z}^{k+1}) = [(2\pi)^n |P_{k+1}|]^{-1/2} \exp -\frac{1}{2} (\underline{x}_{k+1} - \hat{\underline{x}}_{k+1})^T P_{k+1}^{-1} (\underline{x}_{k+1} - \hat{\underline{x}}_{k+1}) \quad (\text{XV})$$

where

$$\hat{\underline{x}}_{k+1} = \underline{f}_{k+1}(\underline{x}_k^*) + \delta \hat{\underline{x}}_{k+1}$$

$$\delta \hat{\underline{x}}_{k+1} = P_{k+1} [H_{k+1}^T R_{k+1}^{-1} \delta \underline{z}_{k+1} + (Q_k^{-1} F_k^T + \frac{1}{2} E_k^T) \Pi_{k+1}^{-1} P_k^{-1} \delta \hat{\underline{x}}_k]$$

$$P_{k+1}^{-1} = (Q_k + F_k P_k F_k^T)^{-1} - 2 E_k^T \Pi_{k+1}^{-1} F_k Q_k^{-1} \\ - E_k^T \Pi_{k+1}^{-1} E_k + H_{k+1}^T R_{k+1}^{-1} H_{k+1} - \sum_{i=1}^m J_{k+1}^i y_{k+1}^i]$$

At  $t_0$  the density is gaussian with mean  $\hat{x}_0$  and covariance  $P_0$ .

$$\hat{x}_0 = x_0^* + \delta \hat{x}_0$$

$$\delta \hat{x}_0 = P_0 [H_0^T R_0^{-1} \delta z_0 + M_0^{-1} \delta a]$$

$$P_0 = [H_0^T R_0^{-1} H_0 - \sum_{i=1}^m J_0^i y_0^i + M_0^{-1}]^{-1}$$

Most of the quantities appearing in (XV) have already been defined. However, the  $y_{k+1}^i$  are components of the vector

$$y_{k+1} \stackrel{\text{Df}}{=} R_{k+1}^{-1} \delta z_{k+1}$$

and

$$E_k \stackrel{\text{Df}}{=} \sum_{i=1}^n G_k^i \Pi_{k+1}^{-1} P_k^{-1} \delta \hat{x}_k q_k^{iT}$$

$$\Pi_{k+1} \stackrel{\text{Df}}{=} F_k^T Q_k^{-1} F_k + P_k^{-1}$$

The  $q_k^{iT}$  is the  $i^{\text{th}}$  row of  $Q_k^{-1}$ .

As was shown in Section 2.1, the minimum mean-square estimate is provided by  $\hat{x}_{k+1}$ . Note that  $P_{k+1}$  is a covariance matrix, so it must always be non-negative definite and preferably should be positive-definite. The manner in which the second order terms enter the defining relation suggests that  $P_{k+1}$  might lose this sign-definiteness if the magnitude of these terms becomes too large. This can provide a criteria for judging when the nominal no longer



provides a reference that is adequate for the gaussian property to remain valid.

Proof: In the subsequent pages, only a detailed outline of the proof is given.

The complete derivation is found in Appendix A.

The a priori statistics for  $\underline{x}_0$ ,  $\{\underline{w}_1\}$ , and  $\{\underline{v}_1\}$  are gaussian and described by (3.1) - (3.3). The desired recursion relations are obtained inductively. First, the initial density function  $p(\underline{x}_0/\underline{z}_0)$  is determined, and then  $p(\underline{x}_{k+1}/\underline{z}^{k+1})$  is derived after assuming the gaussian form for  $p(\underline{x}_k/\underline{z}^k)$ .

$$p(\underline{z}_0/\underline{x}_0) = k_v \exp - \frac{1}{2} \{(\underline{z}_0 - \underline{h}_0)^T R_0^{-1} (\underline{z}_0 - \underline{h}_0)\} \quad (4.1)$$

Expand  $\underline{h}_0$  in a Taylor series about  $\underline{x}_0^*$  and retain only the quadratic terms in  $\delta \underline{x}_0$  that appear in  $(\underline{z}_0 - \underline{h}_0)^T R_0^{-1} (\underline{z}_0 - \underline{h}_0)$ . Then

$$\begin{aligned} p(\underline{z}_0/\underline{x}_0) = k_v \exp - \frac{1}{2} \{ & \delta \underline{z}_0^T R_0^{-1} \delta \underline{z}_0 - 2 \delta \underline{z}_0^T R_0^{-1} H_0 \delta \underline{x}_0 \\ & + \delta \underline{x}_0^T [H_0^T R_0^{-1} H_0 - \sum_{i=1}^m J_{o^i}^T y_{o^i}^i] \delta \underline{x}_0 \} \end{aligned} \quad (4.2)$$

where the notation has been defined above. The nature of the approximation of  $p(\underline{z}_0/\underline{x}_0)$  (and, more generally, the  $p(\underline{z}_k/\underline{x}_k)$ ) has been found to a critical concern in attempting to describe the  $p(\underline{x}_k/\underline{z}^k)$ . This aspect will be discussed in Chapter 6.

Substitution of (3.3) and (4.2) into (2.15) yields

$$p(\underline{x}_0/\underline{z}_0) = k_o \exp - \frac{1}{2} \{(\underline{x}_0 - \hat{\underline{x}}_0)^T P_o^{-1} (\underline{x}_0 - \hat{\underline{x}}_0)\} \quad (4.3)$$

where

$$\hat{\underline{x}}_0 = \underline{x}_0^* + \delta \underline{x}_0$$

$$\delta \hat{\underline{x}}_0 = P_0 [H_0^T R_0^{-1} \delta \underline{z}_0 + M_0^{-1} \delta \underline{a}]$$

$$P_0 = [H_0^T R_0^{-1} H_0 - \sum_{i=1}^m J_{00}^i y_0^i + M_0^{-1}]^{-1}$$

and

$$\delta \underline{a} \stackrel{\text{Df}}{=} \underline{a} - \underline{x}_0^*$$

The minimum mean-square estimate of  $\underline{x}_0$  given the data  $\underline{z}_0$  is  $\hat{\underline{x}}_0$ .

A posteriori density  $p(\underline{x}_{k+1}/\underline{z}^{k+1})$

To determine the relations for an arbitrary sampling time, assume at  $t_k$  that

$$p(\underline{x}_k/\underline{z}^k) = k_k \exp - \frac{1}{2} \{(\underline{x}_k - \hat{\underline{x}}_k)^T P_k^{-1} (\underline{x}_k - \hat{\underline{x}}_k)\} \quad (4.4)$$

The derivation of  $p(\underline{x}_{k+1}/\underline{z}^{k+1})$  is accomplished according to the following steps.

1. Form  $p(\underline{x}_{k+1}/\underline{z}^k)$ .

From (3.1) and (I-N), it is clear that

$$P(\underline{x}_{k+1}/\underline{x}_k) = k_w \exp - \frac{1}{2} \{(\underline{x}_{k+1} - \underline{f}_{k+1})^T Q_k (\underline{x}_{k+1} - \underline{f}_{k+1})\} \quad (4.5)$$

Expand  $\underline{f}_{k+1}$  in a Taylor series and retain only the quadratic terms (and lower order) in  $\delta \underline{x}_k$  in the exponent. This result in combination with (4.4) produces

$$p(\underline{x}_{k+1}/\underline{z}^k) = \frac{k_k k_w}{k_{Nk}} \exp - \frac{1}{2} \{ \delta \underline{x}_{k+1}^T Q_k^{-1} \delta \underline{x}_{k+1} + \delta \hat{\underline{x}}_k^T P_k^{-1} \delta \hat{\underline{x}}_k$$

$$- \delta \underline{v}_k^T B_k^{-1} \delta \underline{v}_k \} \quad (4.6)$$

where

$$\delta \underline{v}_k = B_k [F_k^T Q_k^{-1} \delta \underline{x}_{k+1} + P_k^{-1} \delta \hat{\underline{x}}_k]$$

$$B_k^{-1} = F_k^T Q_k^{-1} F_k - \sum_{i=1}^n G_k^i \omega_k^i + P_k^{-1}$$

and

$$\omega_k \stackrel{\text{Def}}{=} Q_k^{-1} \delta \underline{x}_{k+1}$$

2. Modify  $\delta \underline{v}_k^T B_k^{-1} \delta \underline{v}_k$

The  $\delta \underline{v}_k^T B_k^{-1} \delta \underline{v}_k$  is not a quadratic function of  $\delta \underline{x}_{k+1}$ , so it must be modified in order for the  $p(\underline{x}_{k+1}/\underline{z}^{k+1})$  to have the gaussian form. Using a Neumann series and neglecting all terms of order greater than quadratic in  $\delta \underline{x}_{k+1}$ , one obtains

$$\begin{aligned} \delta \underline{v}_k^T B_k^{-1} \delta \underline{v}_k &= \delta \hat{\underline{x}}_k^T P_k^{-1} \Pi_{k+1}^{-1} P_k^{-1} \delta \hat{\underline{x}}_k + 2 \delta \hat{\underline{x}}_k^T P_k^{-1} \Pi_{k+1}^{-1} [F_k^T Q_k^{-1} + \frac{1}{2} E_k] \delta \underline{x}_{k+1} \\ &\quad + \delta \underline{x}_{k+1}^T [Q_k^{-1} F_k \Pi_{k+1}^{-1} F_k^T Q_k^{-1} + 2 E_k^T \Pi_{k+1}^{-1} F_k Q_k^{-1} \\ &\quad + E_k^T \Pi_{k+1}^{-1} E_k] \delta \underline{x}_{k+1} \end{aligned} \quad (4.7)$$

where

$$\begin{aligned} \Pi_{k+1} &= F_k^T Q_k^{-1} F_k + P_k^{-1} \\ E_k &= \sum_{i=1}^n G_k^i \Pi_{k+1}^{-1} P_k^{-1} \delta \hat{\underline{x}}_k \mathbf{q}_k^i{}^T \end{aligned}$$

The  $\mathbf{q}_k^i{}^T$  is the  $i^{\text{th}}$  row of  $Q_k^{-1}$ .

3. Determine  $p(\underline{z}_{k+1}/\underline{x}_{k+1})$

From (II-N) and (4.1), one sees that

$$p(\underline{z}_{k+1}/\underline{x}_{k+1}) = k_v \exp - \frac{1}{2} \{ (\underline{z}_{k+1} - \underline{h}_{k+1})^T R_{k+1}^{-1} (\underline{z}_{k+1} - \underline{h}_{k+1}) \} \quad (4.8)$$

Expand  $h_{k+1}$  in a Taylor series and retain only the appropriate terms. Then

$$\begin{aligned}
 p(\underline{z}_{k+1}/\underline{x}_{k+1}) &= k_v \exp - \frac{1}{2} \{ \delta \underline{z}_{k+1}^T R_{k+1}^{-1} \delta \underline{z}_{k+1} \} \\
 &\quad \exp - \frac{1}{2} \{ \delta \underline{x}_{k+1}^T [H_{k+1}^T R_{k+1}^{-1} H_{k+1} - \sum_{i=1}^m J_{k+1}^i y_{k+1}^i] \\
 &\quad \delta \underline{x}_{k+1} - 2 \delta \underline{z}_{k+1}^T R_{k+1}^{-1} H_{k+1} \delta \underline{x}_{k+1} \} \quad (4.9)
 \end{aligned}$$

4. Form  $p(\underline{z}_{k+1}/\underline{z}^k)$

It follows in a straightforward manner from (IV), (4.6), (4.7), and (4.9) that

$$\begin{aligned}
 p(\underline{z}_{k+1}/\underline{z}^k) &= \frac{k_k^k w^k v}{k_{k+1}^k N_k} \exp - \frac{1}{2} \{ \delta \underline{z}_{k+1}^T R_{k+1} \delta \underline{z}_{k+1} - \delta \hat{\underline{x}}_{k+1}^T P_{k+1}^{-1} \delta \hat{\underline{x}}_{k+1} \\
 &\quad + \delta \hat{\underline{x}}_k^T [P_k^{-1} - P_k^{-1} \Pi_{k+1}^{-1} P_k^{-1}] \delta \hat{\underline{x}}_k \} \quad (4.10)
 \end{aligned}$$

where  $\delta \hat{\underline{x}}_{k+1}$  and  $P_{k+1}$  shall be defined below.

5. Form  $p(\underline{x}_{k+1}/\underline{z}^{k+1})$

Performing the operations indicated by (IV), the a posteriori density is found to be given by (XV) thereby completing the proof.

Q. E. D.

## 4.2 ON THE APPROXIMATION OF NONLINEAR SYSTEMS

Commonly, the analysis of the nonlinear system (I-N) and (II-N) is approached by introducing linear perturbation theory. This requires the choice of nominal values  $\underline{x}_{k-1}^*$  for the state. Then, the linear perturbation equations are

$$\delta \underline{x}_k = F_k \delta \underline{x}_{k-1} + \underline{w}_{k-1} \quad (4.11)$$

$$\delta \underline{z}_k = H_k \delta \underline{x}_k + \underline{v}_k \quad (4.12)$$

where the  $F_k$  and  $H_k$  have been defined in Section 4.1. The choice of the nominal is made somewhat arbitrarily, and its adequacy is gauged by the subsequent results.

Assuming that (4.11) and (4.12) are accurate representations of the deviations from the nominal, the problem of estimating  $\underline{x}_k$  reduces to the simpler problem of estimating  $\delta \underline{x}_k$ . Since this system is linear and the noise sequences are gaussian, the recursive minimum mean-square estimate of  $\delta \underline{x}_k$  is given by the Kalman filter equations of Section 3.1. The result of the preceding section gives a generalization of this linear approximation, and in so doing, provides insight into other aspects of the problem, including the choice of the nominal.

#### 4.2.1 Relation to the Kalman Filter

In this section we shall demonstrate that (XV) reduces to the Kalman case when the matrices containing the second partial derivatives are identically zero. Let

$$G_k^i = J_{k+1}^i = 0 \quad \text{for all } i, j$$

Then, it is true that

$$E_k \equiv 0.$$

The equation for  $\delta \hat{\underline{x}}_{k+1}$  in (XV) reduces to

$$\delta \hat{\underline{x}}_{k+1} = P_{k+1} [H_{k+1}^T R_{k+1}^{-1} \delta \underline{z}_{k+1} + Q_k^{-1} F_k \Pi_{k+1}^{-1} P_k^{-1} \delta \hat{\underline{x}}_k]$$

and the covariance  $P_{k+1}$  becomes

$$P_{k+1}^{-1} = [(Q_k + F_k P_k F_k^T)^{-1} + H_{k+1}^T R_{k+1}^{-1} H_{k+1}]$$

Let

$$P'_{k+1} \stackrel{\text{Def}}{=} Q_k + F_k P_k F_k^T \quad (4.13)$$

From a matrix inversion lemma [46, 30], it follows that

$$P_{k+1} = P'_{k+1} - K_{k+1} H_{k+1}^T P'_{k+1} \quad (4.14)$$

where

$$K_{k+1} \stackrel{\text{Def}}{=} P'_{k+1} H_{k+1}^T [H_{k+1}^T P'_{k+1} H_{k+1} + R_{k+1}]^{-1} \quad (4.15)$$

But (4.13) - (4.15) correspond to the gain and error covariance matrix of the Kalman filter equation (XI) with  $F_k$  substituted for  $\hat{\Phi}_{k+1,k}$ .

The estimate  $\delta \hat{x}_{k+1}$  can be modified since it is known [30] that (4.15) can be written as

$$K_{k+1} = P_{k+1} H_{k+1}^T R_{k+1}^{-1} \quad (4.16)$$

Then

$$\delta \hat{x}_{k+1} = K_{k+1} \delta z_{k+1} + P_{k+1} Q_k^{-1} F_k \Pi_{k+1}^{-1} P_k^{-1} \delta \hat{x}_k$$

Substitute the defining relation for  $\Pi_{k+1}^{-1}$ . Then this becomes

$$\delta \hat{x}_{k+1} = K_{k+1} \delta z_{k+1} + P_{k+1} P_{k+1}^{-1} F_k \delta \hat{x}_k$$

But

$$P_{k+1} P_{k+1}^{-1} = (I - K_{k+1} H_{k+1}^T)$$

so

$$\delta \hat{\underline{x}}_{k+1} = F_k \delta \hat{\underline{x}}_k + K_{k+1} [\delta \underline{z}_{k+1} - H_{k+1} F_k \delta \underline{x}_k]$$

which is the Kalman estimate.

#### 4.2.2 Choice of the Nominal

The values of the state that are chosen as the nominal will obviously play a key role in determining the validity of the approximation. For many problems (e.g., space navigation), it is convenient to specify a nominal before the system is in operation and to then compute many of the quantities required by the filter off-line. This policy minimizes the amount of computation that must be performed while the system is in operation. It has been suggested that it is not always desirable to prespecify the nominal because the quality of the linear approximation is caused to deteriorate more rapidly. The filter described by (XV) provides analytical corroboration of this intuitive idea and demonstrates that the best choice of nominal at each sampling time  $t_{k+1}$  is the  $\hat{\underline{x}}_k$ .

In (XV) suppose that the nominal is selected as

$$\underline{x}_k^* \equiv \hat{\underline{x}}_k$$

with this nominal, it is obvious from

$$\delta \hat{\underline{x}}_k \stackrel{\text{Df}}{=} \hat{\underline{x}}_k - \underline{x}_k^*$$

that

$$\delta \hat{\underline{x}}_k \equiv 0.$$

The estimate (XV) reduces immediately to

$$\delta \hat{\underline{x}}_{k+1} = P_{k+1} H_{k+1} R_{k+1}^{-1} \delta \underline{z}_{k+1} \quad (4.17)$$

Furthermore, from the definition of  $E_k$ , observe that

$$E_k \equiv 0.$$

Therefore,

$$P_{k+1} = [(P'_{k+1})^{-1} + H_{k+1}^T R_{k+1}^{-1} H_{k+1} - \sum_{i=1}^m J_{k+1}^i y_{k+1}^i]^{-1} \quad (4.18)$$

Certainly (4.17) and (4.18) are simpler in appearance than their counterparts in (XV). In fact, one notices that the second order term  $G_k^i$  from the plant has disappeared entirely. The error covariance matrix  $P_{k+1}$  contains the second order measurement effects  $J_{k+1}^i$ , and these terms cause the  $P_{k+1}$  to depend upon the measurement data. This is in sharp contrast with the Kalman filter in which the error covariance matrix and, therefore, the gain can be computed off-line.

With this choice of nominal, the minimum mean-square estimate is seen to be

$$\hat{x}_{k+1} = f_{k+1}(\hat{x}_k) + K_{k+1} [z_{k+1} - h_{k+1}(\xi_{k+1})] \quad (4.19)$$

where  $K_{k+1}$  is defined by (4.16) and

$$\xi_{k+1} = f_{k+1}(\hat{x}_k)$$

The  $P_{k+1}$  in (4.18) can be written in a form that is more computationally attractive. Let

$$P_{k+1}^L \stackrel{\text{Df}}{=} [(P'_{k+1})^{-1} + H_{k+1}^T R_{k+1}^{-1} H_{k+1}]^{-1}$$

By a matrix inversion lemma, this is equal to

$$P_{k+1}^L = P'_{k+1} - P'_{k+1} H_{k+1}^T (H_{k+1} P'_{k+1} H_{k+1}^T + R_{k+1}^{-1})^{-1} H_{k+1} P'_{k+1}$$



With this definition, (4.18) becomes

$$P_{k+1}^L = P_{k+1}^L \left[ I - \left( \sum_{i=1}^m J_{k+1}^i y_{k+1}^i \right) P_{k+1}^L \right]^{-1}$$

When the second order terms are not present, this becomes

$$P_{k+1} = P_{k+1}^L$$

and we note that  $P_{k+1}^L$  is the error covariance matrix of the Kalman filter.

#### 4.2.3 Conclusions

The preceding development has produced several interesting results.

- 1) Linearization of the nonlinear plant and measurement equations about some nominal does not provide the most general form for the mean and covariance of a gaussian conditional probability density function  $p(\underline{x}_k / \underline{z}^k)$ .
- 2) Expansion of the nonlinear plant and measurement equations about arbitrary nominal values subject to the constraint that the density  $p(\underline{x}_k / \underline{z}^k)$  must be gaussian produces mean and covariance that depend upon the second order terms of the expansions. In contrast with the Kalman filter, the covariance depends upon the measurement data.
- 3) If for each  $k$  the plant equation  $\underline{f}_k$  is expanded about the conditional mean  $\hat{\underline{x}}_{k-1}$ , all second order terms from the plant equation are eliminated in the relation for  $\hat{\underline{x}}_k$ . The covariance still depends upon the measurement data and contains second order terms of the measurement equation.

- 4) The defining relation for the conditional variance  $P_{k+1}$  contains negative terms that could destroy the positive- (or non-negative-) definiteness property required for this matrix. If such a situation were to arise, it would suggest that the nominal values were no longer an adequate reference and would suggest that a nongaussian conditional density function is required.

The disadvantages inherent in these results arise primarily through the increased number of computations that must be performed on-line. That is, if the minimum mean-square estimate provided by  $\hat{x}_{k+1}$  in (4.20) is used, the error covariance matrix must be computed during the operation of the system described by (I-N) and (II-N). In the Kalman filter, this matrix does not depend upon the measurements, so it can be computed in advance if the nominal has been specified. Also, if the estimate is treated as the nominal, it is necessary to compute all system matrices on-line since the  $F_k$ ,  $H_k$ , and  $J_{k+1}^i$  are all computed using the nominal values. Thus, the computational load is greatly increased if this formulation is to be implemented. Additional remarks regarding the nonlinear filter of Section 4.1 are found in Chapters 5 and 7.

#### 4.3 ON THE CONTROL OF A LINEAR PLANT USING NONLINEAR MEASUREMENT DATA

In this section, no formal results are to be exhibited. Rather, the filter derived in Section 4.1 will be utilized to suggest that the Separation Principle of Section 3.2 cannot be extended to the situation in which the plant is given by

$$\underline{x}_k = \Phi_{k,k-1} \underline{x}_{k-1} + \Gamma_{k,k-1} \underline{u}_{k-1} + \underline{w}_{k-1} \quad (\text{I-L})$$

and the measurements are described by

$$\underline{z}_k = \underline{h}_k(\underline{x}_k) + \underline{v}_k \quad (\text{II-N})$$

As in Chapter 3, consider the problem of establishing the control policy for the system (I-L) - (II-N) that minimizes the quadratic performance index (III<sub>C</sub>-L) under the constraint on  $\underline{h}_k$  that the a posteriori density remain gaussian for all k.

For the last stage, the function  $\bar{J}_N$  is easily seen to be (re: equation 3.40)

$$\begin{aligned} \bar{J}_N = & \underline{x}_{N-1}^T \Phi_{N,N-1}^T W_{N,N-1}^X \Phi_{N,N-1} \underline{x}_{N-1} + 2 \underline{x}_{N-1}^T \Phi_{N,N-1}^T W_{N,N-1}^X \Gamma_{N,N-1}^T \underline{u}_{N-1} \\ & + \underline{u}_{N-1}^T (W_{N,N-1}^U + \Gamma_{N,N-1}^T W_{N,N-1}^X \Gamma_{N,N-1}) \underline{u}_{N-1} + \text{trace} [W_{N,N-1}^X Q_{N-1}] \quad (4.21) \end{aligned}$$

The control that minimizes  $E[\bar{J}_N / \underline{z}^{N-1}]$  is

$$\underline{u}_{N-1}^0 = -\Lambda_N \Phi_{N,N-1} \hat{\underline{x}}_{N-1} \quad (4.22)$$

where  $\Lambda_N$  is defined by (3.42). In this case the estimate is not given by the Kalman filter equations. It follows from Section 4.1 and the fact that the  $\underline{u}^{N-2}$  is known that

$$\begin{aligned} \hat{\underline{x}}_{N-1} = & \Phi_{N-1,N-2} \hat{\underline{x}}_{N-2} + K_{N-1} (\underline{z}_{N-1} - H_{N-1} \Phi_{N-1,N-2} \hat{\underline{x}}_{N-2}) \\ & + (I - K_{N-1} H_{N-1}) \Gamma_{N-1,N-2} \underline{u}_{N-2} \quad (4.23) \end{aligned}$$

where

$$\begin{aligned} K_{N-1} &= P_{N-1} H_{N-1}^T R_{N-1}^{-1} \\ P_{N-1}^{-1} &= P_{N-1}'^{-1} + H_{N-1}^T R_{N-1}^{-1} H_{N-1} - \sum_{i=1}^m J_{N-1}^i y_{N-1}^i \end{aligned}$$

$$P'_{N-1} = \Phi_{N-1, N-2}^T P_{N-2} \Phi_{N-1, N-2}^T + Q_{N-1}$$

Thus, the Separation Principle is valid for a single-stage problem.

Unfortunately, it does not appear to be possible to extend this to multi-stage problems. We shall consider a two-stage problem and indicate the reason for the added difficulty. No attempt will be made to derive the control law for this problem.

As in (3.43) of Section 3.2, the  $\bar{z}_N^0$  can be written as

$$\begin{aligned} \bar{z}_N^0 &= \underline{x}_{N-1}^T \Phi_{N, N-1}^T \Pi_N \Phi_{N, N-1} \underline{x}_{N-1} \\ &\quad + \tilde{\underline{x}}_{N-1}^T \Phi_{N, N-1}^T \Pi_N' \Gamma_{N, N-1} \Lambda_N \Phi_{N, N-1} \tilde{\underline{x}}_{N-1} \\ &\quad + \text{trace} (W_N^X Q_{N-1}) \end{aligned} \quad (4.24)$$

From  $\bar{z}_{N-1}$  according to  $(V_C)$

$$\begin{aligned} \bar{z}_{N-1} &= \int \bar{z}_N^0 \delta(\underline{z}_{N-1} - \underline{h}_{N-1}) p(\underline{v}_{N-1}) d(\underline{v}_{N-1}, \underline{z}_{N-1}) \\ &= \underline{x}_{N-1}^T \Phi_{N, N-1}^T \Pi_N \Phi_{N, N-1} \underline{x}_{N-1} + \text{trace} (W_N^X Q_{N-1}) \\ &\quad + \int \tilde{\underline{x}}_{N-1}^T \Phi_{N, N-1}^T \Pi_N' \Gamma_{N, N-1} \Lambda_N \Phi_{N, N-1} \tilde{\underline{x}}_{N-1} \\ &\quad \delta(\underline{z}_{N-1} - \underline{h}_{N-1}) p(\underline{v}_{N-1}) d(\underline{v}_{N-1}, \underline{z}_{N-1}) \end{aligned} \quad (4.25)$$

In contrast with the linear filter, the term involving  $\tilde{\underline{x}}_{N-1}$  does contribute to the control. To verify this, we shall perform the integration of (4.25) with respect to  $\underline{z}_{N-1}$ . Then, as indicated in the formation of  $\bar{z}_{N-1}$  in  $(V_C)$ , one must integrate with respect to  $\underline{x}_{N-1}$ . After these two integrations are performed, the error can be written as

$$\begin{aligned} \tilde{\underline{x}}_{N-1} = & \Phi_{N-1, N-2} \tilde{\underline{x}}_{N-2} + K_{N-1} [\underline{h}_{N-1} + \underline{v}_{N-1} - H_{N-1} \\ & (\Phi_{N-1, N-2} \hat{\underline{x}}_{N-2} + \Gamma_{N-1, N-2} \underline{u}_{N-2})] - \underline{w}_{N-2} \end{aligned} \quad (4.26)$$

As has been indicated in Section (4.1), the  $\underline{h}_N$  is approximately

$$\underline{h}_{N-1} = H_{N-1} \underline{x}_{N-1} + \nabla^2 \underline{h} \quad (4.27)$$

where

$$\nabla^2 \underline{h} \stackrel{\text{Df}}{=} \begin{bmatrix} \underline{x}_{N-1}^T J_{N-1}^1 \underline{x}_{N-1} \\ \vdots \\ \underline{x}_{N-1}^T J_{N-1}^m \underline{x}_{N-1} \end{bmatrix}$$

Introducing (4.27) into (4.26), one obtains

$$\begin{aligned} \tilde{\underline{x}}_{N-1} = & \Phi_{N-1, N-2} \tilde{\underline{x}}_{N-2} + K_{N-1} [H_{N-1} \underline{x}_{N-1} + \nabla^2 \underline{h} + \underline{v}_{N-1} \\ & - H_{N-1} (\Phi_{N-1, N-2} \hat{\underline{x}}_{N-2} + \Gamma_{N-1, N-2} \underline{u}_{N-2})] - \underline{w}_{N-2} \end{aligned}$$

But  $\underline{x}_{N-1}$  is described by (I-L) so this becomes

$$\tilde{\underline{x}}_{N-1} = (I - K_{N-1} H_{N-1}) (\Phi_{N-1, N-2} \hat{\underline{x}}_{N-2} - \underline{w}_{N-1}) + K_{N-1} \nabla^2 \underline{h} + K_{N-1} \underline{v}_{N-1}$$

At this juncture, let us recall that in the linear problem, the  $\nabla^2 \underline{h}$  would be identically zero and the gain  $K_{N-1}$  is independent of the measurements. Thus, as was stated in Chapter 3, the term involving  $\tilde{\underline{x}}_{N-1}$  does not contribute to the control policy. It is this fact that permits the proof of the Separation Principle for linear systems. It is clear from (4.15) and (XV) that  $K_{N-1}$  depends upon the measurement data  $\underline{z}_{N-1}$  and must, therefore, contain the

control  $\underline{u}_{N-2}$ . Furthermore, the  $\nabla^2 \underline{h}$  depends upon  $\underline{x}_{N-1}$ , so from (I-L), it must be true that  $\nabla^2 \underline{h}$  must contain  $\underline{u}_{N-2}$ .

The  $\nabla^2 \underline{h}$  and the  $K_{N-1}$  do not allow the term involving  $\tilde{\underline{x}}_{N-1}$  in (4.25) to be neglected in determining the control policy. This was necessary in establishing the Separation Principle in Section 3.2. Of course, if one were to continue the derivation, it might be found that the control is unaffected by these terms, but this would be surprising.

## CHAPTER FIVE

### FILTERING FOR NAVIGATION OF A SPACECRAFT

Several of the results of Chapters 3 and 4 are applied to the problem of estimating the state of a spacecraft. The linear filter of Chapter 3 is utilized for the greater part of the study contained in this chapter. In many cases, linear perturbation theory is found to adequately describe the physical system, so one would expect the linear filter to perform "satisfactorily". Occasions do arise, however, when the nonlinear effects seriously affect and sometimes even destroy the validity of the output of the linear filter. It is the intent in this chapter to illustrate both of these situations. Then, several techniques are investigated which allow the range of applicability of linear theory to be considerably extended. Finally, the nonlinear filter of Chapter 4 is applied to the problem to illustrate the effect of including nonlinear terms. These results, unfortunately, are of a somewhat disappointing nature.

The basic problem and the mathematical model are discussed in Section 5.1. The numerical results obtained from the digital computer simulation of the problem are presented in Section 5.2. The conclusions that can be drawn from these results are presented in Section 5.3.

#### 5.1 THE SPACE NAVIGATION PROBLEM

The objective in this chapter is to consider the applicability of perturbative techniques to a significant nonlinear problem. In particular, the problem of estimating the position and velocity (i. e. , the state) of a spacecraft moving in a nearly circular orbit about the Earth is studied. The estimates are to be based upon the measurements provided by a horizon sensor aboard the craft.

Several different estimation policies are utilized, although the linear filter described by (XI) provides the basic configuration. The policies are listed and described below. Before discussing them, the basic mathematical model shall be presented. A more detailed discussion can be found in Appendix B.

#### 5.1.1 The Mathematical Model

For this study, the Earth shall be assumed to be spherical with radius  $r_0$  and to have a spherical gravity potential  $U$  described by

$$U = \frac{\mu}{R}$$

The  $\mu$  is a constant equal to the product of the mass of the Earth and the universal gravitational constant. Let  $R$  be the distance from the center of the Earth to the spacecraft.

A coordinate system is defined to be a nonrotating cartesian system with origin at the center of the Earth. The coordinate axes shall be denoted by  $X$ ,  $Y$ ,  $Z$ . The motion shall be assumed to occur, primarily, in the  $X$ - $Y$  plane. In this system, the equations of motion for the spacecraft are known to be

$$\ddot{\underline{R}} = - \frac{\mu}{R^3} \underline{R} \quad (5.1)$$

In order to use state vector notation, (5.1) must be reduced to a first order differential equation. This is accomplished by defining the state  $\underline{x}$  to be the six-dimensional vector formed from the components of the position  $\underline{R}$  and velocity  $\underline{V}$  vectors.



$$\underline{x} \stackrel{\text{Df}}{=} \begin{bmatrix} \underline{R} \\ \underline{V} \end{bmatrix} \stackrel{\text{Df}}{=} \begin{bmatrix} X \\ Y \\ Z \\ \dot{X} \\ \dot{Y} \\ \dot{Z} \end{bmatrix} \equiv \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \\ X_6 \end{bmatrix}$$

Then, the state is seen to evolve according to

$$\dot{\underline{x}} = \begin{bmatrix} \dot{\underline{R}} \\ \dot{\underline{V}} \end{bmatrix} = \begin{bmatrix} \underline{V} \\ -\frac{\mu}{R^3} \underline{R} \end{bmatrix} \stackrel{\text{Df}}{=} \underline{f}(\underline{x}) \quad (5.2)$$

For this system we shall assume that the plant does not contain any noise, so (5.2) provides the specific form for the plant equation (I) to be considered in this example.

The position and velocity of the spacecraft are to be estimated using the angular measurements from a horizon sensor. This instrument is assumed to measure:

- (1) the direction of the local vertical relative to the X-axis of the coordinate system. The direction is specified by the two angles  $\alpha$  and  $\delta$ , where  $\alpha$  is the angle between the X-Y plane and the line of sight and  $\delta$  is the angle between the X-axis and the projection of the line of sight onto the X-Y plane.
- (2) the subtended Earth angle  $\beta$ . The  $\beta$  is defined as the angle between the line of sight to the edge of the planet and the local vertical.

These angles are depicted in Figure B-1 of Appendix B and are given by

$$\left. \begin{aligned} \alpha &= -\sin^{-1} \frac{X_3}{R} \\ \delta &= \sin^{-1} \frac{X_2}{(X_1^2 + X_2^2)^{1/2}} \\ \beta &= \sin^{-1} \frac{r_o}{R} \end{aligned} \right\} \quad (5.3)$$

Let

$$\underline{h}(\underline{x}) \stackrel{\text{Df}}{=} \begin{bmatrix} \alpha \\ \delta \\ \beta \end{bmatrix} \quad (5.4)$$

Assume the measurements contain an additive, gaussian white noise sequence.

Then, (5.4) completes the definition of the measurement equation (II-N) for this example.

The nonlinear equations (5.2) and (5.4) must be expanded in Taylor series relative to some choice of nominal values for the state. To apply the Kalman filter of Chapter 3, the system must be reduced to a linear model, whereas second order terms are required for the filter of Chapter 4. Assume that the required nominal  $\underline{x}^*$  exists and expand (5.2) and (5.3) in a Taylor series. In this chapter the plant equation will always be assumed to be linear, so one gets

$$\delta \dot{\underline{x}} = F \delta \underline{x} \quad (5.5)$$

where  $F$  is the matrix containing the partial derivatives of  $\underline{f}$  with respect to  $\underline{x}$ . Let

$$\delta \underline{x} \stackrel{\text{Df}}{=} \underline{x} - \underline{x}^*$$

The solution of (5.5) is known to have the form [3]

$$\delta \underline{x}_k = \Phi_{k,k-1} \delta \underline{x}_{k-1} \quad (5.6)$$

where  $\Phi_{k,k-1}$  is the state transition matrix and is the solution of

$$\dot{\Phi} = F \Phi$$

with initial condition

$$\Phi(t_0, t_0) = I$$

For the dynamical system (5.2), it is possible to obtain  $\Phi$  in a closed form [55].

The solution is presented in Appendix B. Equation (5.6) will serve as the plant equation for the perturbed state. Note again that no noise appears in this relation.

The first and second order partial derivatives of  $\alpha$ ,  $\delta$ , and  $\beta$  are formed in a straightforward manner. They are presented in Appendix B. In many instances, the partial derivatives are very difficult to determine analytically because of the complicated nature of the equations. Wengert [47] has suggested a procedure for determining these derivatives in terms of elementary functions that seems to be quite reasonable. Wilkins [48] applied this approach to a complicated system and concluded that the method was very satisfactory. We mention this work because it appears to be a necessary consideration for the development of a practical nonlinear perturbation theory.

#### 5.1.2 The Estimation Policies

Most of the results presented in the next section are based upon the linear filter described by Lemma 3.1 of Chapter 3. Five different policies are examined using the linear filter. Two additional policies are investigated using

the nonlinear filter described by Lemma 4.1 of Chapter 4. Each of the policies is discussed in the detail deemed necessary in the succeeding paragraphs.

(L-1) Linear filter with a prespecified nominal

The linearization is performed relative to a circular orbit at a 100 n. mile altitude. This nominal is used throughout the flight and the filter equations (XI) are utilized.

(L-2) Linear filter using  $\hat{\underline{x}}_{k-1}$  as the nominal state at each  $t_k$

It was observed in Section 4.2 that the most appropriate choice of nominal at each sampling time  $t_k$  is the estimate  $\hat{\underline{x}}_{k-1}$ . Thus, at every sampling time, the nominal is selected to be

$$\underline{x}_{k-1}^* = \hat{\underline{x}}_{k-1} = \underline{f}_{k-1}(\hat{\underline{x}}_{k-2}) + \delta \hat{\underline{x}}_{k-1}$$

The linearization is accomplished relative to this nominal, and (XI) is again utilized. In this case, observe that

$$\delta \hat{\underline{x}}_{k-1} = 0$$

after the change of nominal has been completed. This procedure shall be referred to as rectification.

Rectification has a disadvantage in that all of the system matrices (i. e.,  $\Phi_{k,k-1}$ ,  $H_k$ , etc.) must be recomputed at each sampling time. If a prespecified nominal is used, the system can be computed and stored prior to the actual realization of the system. In fact, the error covariance matrix  $P_k$  and the gain matrix  $K_k$  can also be pre-computed. This results in a considerable reduction of on-line computation. On the other hand, the additional on-line computation

does not represent a significant restriction in many cases, so rectification provides a very sensible means of extending linear theory.

Two means of extending the linear theory without resorting to orbit rectification are suggested. The first can be applied to systems containing plant noise as well as measurement noise. The second policy is restricted to systems with noise-free plants. More details regarding these policies follow immediately.

(L-3) Modified Observation Matrix  $H_k$

The approach used in establishing this policy shall be discussed in somewhat greater detail in Chapter 6. For the moment, consider the measurements to be described by (II-N).

$$\underline{z}_k = \underline{h}_k(\underline{x}_k) + \underline{v}_k \quad (\text{II-N})$$

Assuming a nominal, expand  $\underline{h}_k$  in a Taylor series and retain the first and second order terms. The  $i^{\text{th}}$  component is given by

$$\begin{aligned} \delta z_k^i &\stackrel{\text{Df}}{=} z_k^i - h_k^i(\underline{x}_k^*) \\ &= \sum_{j=1}^n \frac{\partial h_k^i}{\partial x_k^j} \delta x_k^j + \sum_{j=1}^n \sum_{\ell=1}^n \frac{\partial^2 h_k^i}{\partial x_k^j \partial x_k^\ell} \delta x_k^j \delta x_k^\ell + v_k^i \quad (5.7) \end{aligned}$$

As has been noted, the  $i^{\text{th}}$  row of the observation matrix  $H_k$  is composed of the partial derivatives  $\partial h_k^i / \partial x_k^j$ . (Naturally, all of the partial derivatives are evaluated with the nominal values  $\underline{x}_k^*$ .)

In approximating  $\underline{h}_k$  by  $\underline{H}_k \delta \underline{x}_k$ , it is clear that the second order (and higher) terms are neglected. But this can be circumvented to a certain extent through the following artifice. Consider the predicted estimate of  $\delta \underline{x}_k$  (i.e.,  $E[\delta \underline{x}_k / \delta \underline{z}^{k-1}]$ ). This estimate depends upon  $\delta \hat{\underline{x}}_{k-1}$  and the plant equations. If the plant is linear, then

$$\delta \hat{\underline{x}}_k' = \Phi_{k,k-1} \delta \hat{\underline{x}}_{k-1}$$

As long as the error in this estimate is small compared to the estimate itself, (5.7) can be approximated by

$$\delta z_k^i \approx \sum_{j=1}^n \left[ \frac{\partial h_k^i}{\partial x_k^j} + \sum_{\ell=1}^n \frac{\partial^2 h_k^i}{\partial x_k^j \partial x_k^\ell} \delta \hat{x}_k^{\ell'} \right] \delta x_k^j + v_k^i \quad (5.8)$$

In (5.8), the  $\delta \hat{\underline{x}}_k'$  has been substituted for the state perturbations  $\delta \underline{x}_k$ . The relationship between the measurements and the state is linear, but it contains the second order partial derivatives. The elements of the observation matrix are redefined as

$$H_k^{ij} = \frac{\partial h_k^i}{\partial x_k^j} + \sum_{\ell=1}^n \frac{\partial^2 h_k^i}{\partial x_k^\ell \partial x_k^j} \delta \hat{x}_k^{\ell'} \quad (5.9)$$

This modified observation matrix is used in conjunction with the linear filter and prespecified nominal described in Case (L-1), and constitutes the third policy. Note that if rectification is utilized, the predicted estimate is identically zero and  $\underline{H}_k$  is reduced to the first order terms.

(L-4) Re-estimation of the initial state

Since there is no noise in the plant, a smoothed estimate of the initial state deviation can be easily determined from  $\delta \hat{\underline{x}}_k$ . In particular, assuming that

$$\delta \underline{x}_k = \Phi_{k,0} \delta \underline{x}_0$$

it follows that

$$\begin{aligned} \delta \hat{\underline{x}}_{0/k} &= E[\delta \underline{x}_0 / \delta \underline{z}^k] \\ &= \Phi_{0,k} \delta \hat{\underline{x}}_k \end{aligned} \quad (5.10)$$

where  $\delta \hat{\underline{x}}_k$  is the estimate provided by the linear filter equations.

Using the smoothed estimate  $\delta \hat{\underline{x}}_{0/k}$  and the true equations of motion, one can determine an estimate of the current deviation that eliminates the errors that accrue through the linear approximation of the plant behavior. To determine this estimate, let  $\hat{\underline{x}}_{0/k}$  be the estimate of the initial state

$$\hat{\underline{x}}_{0/k} = \underline{x}_0^* + \delta \hat{\underline{x}}_{0/k}$$

Then, the estimate of the current state can be computed as

$$\hat{\underline{x}}_k = f_k(\hat{\underline{x}}_{0/k})$$

The  $\hat{\underline{x}}_k$  needs to be computed only when it is actually required

(e.g., when a guidance maneuver is to be introduced that is based

upon  $\hat{\underline{x}}_k$  or  $\delta \hat{\underline{x}}_k = \hat{\underline{x}}_k - \underline{x}_k^*$ ). For this policy the nominal is pre-specified.

(L-5) Rectification using a smoothed estimate

It is possible at each sampling time  $t_k$  to form a smoothed estimate  $\delta \hat{\underline{x}}_{k-1/k}$  using (3-28). One would expect that this estimate would generally be superior to  $\delta \hat{\underline{x}}_{k-1}$  since it is based upon more data. Then, let the nominal state be selected as

$$\underline{x}_{k-1}^* = (\underline{x}_{k-1}^*)^- + \delta \hat{\underline{x}}_{k-1/k}$$

The old nominal  $(\underline{x}_{k-1}^*)^-$  is computed from  $\delta \hat{\underline{x}}_{k-1}$  as discussed in Case (L-2). The smoothed estimate is only used to modify the nominal so the estimate of the perturbation  $\delta \hat{\underline{x}}_{k-1}$  used in computing  $\delta \hat{\underline{x}}_k$  is set equal to  $-\delta \hat{\underline{x}}_{k-1/k}$ . This policy essentially doubles the computational requirements required in Case (L-2).

This completes the definition of the estimation policies that are investigated using the linear filter. Cases (L-1) and (L-2) are repeated using the nonlinear filter described in Chapter 4.

(N-1) Nonlinear filter with a prespecified nominal

(N-2) Nonlinear filter using  $\hat{\underline{x}}_{k-1}$  as the nominal state at each  $t_k$

These policies describe the basic nature of the numerical investigation. These cases were investigated for a variety of sampling intervals, instrument accuracies, deviations in initial conditions, and random noise sequences. The data that are presented in Section 4.2 are representative of the type of results that were obtained. In order to most clearly demonstrate the character of the results, a minimum amount of data has been included.



Two types of data are presented. Certainly, the error covariance matrix  $P_k$  should describe the effectiveness of the filtering procedure if the model is accurate. Since we are approximating a nonlinear system, the  $P_k$  does not always reflect the covariance of the error in the estimate, so a Monte Carlo [61] simulation (i.e., a random number generator is used to simulate the noise in the measurements) is performed to obtain samples of the actual error in the estimate. These errors are compared with the semi-axes of the position and velocity error ellipsoids [10, 30] in an effort to determine if the error covariance matrix is a valid measure of the errors.

## 5.2 NUMERICAL RESULTS

The trajectory that is examined is approximately a circular orbit at 100 n.miles altitude. The constants assumed for the Earth model are [63]

$$\mu = 1.4076539 \times 10^{16} \text{ ft}^3/\text{sec}^2$$

$$r_o = 20,925,738.0 \text{ ft.}$$

The initial conditions for the nominal (i.e., before any measurements) are designed to give a 100 n. mile circular orbit based on these constants.

$$\underline{R} = \begin{bmatrix} 21,533,349.7 \\ 0 \\ 0 \end{bmatrix} ; \quad \underline{V} = \begin{bmatrix} 0 \\ 25,567.728 \\ 0 \end{bmatrix}$$

The initial conditions for the actual trajectory are unknown, but the deviation from the nominal is assumed to belong to a gaussian ensemble with mean zero and prescribed covariance matrix  $M_o$ . The  $M_o$  is assumed to be diagonal with the general form

$$M_o = \begin{bmatrix} \sigma_p^2 I & 0 \\ 0 & \sigma_v^2 I \end{bmatrix}$$

where  $\sigma_p^2$  and  $\sigma_v^2$  are the variances of the position and velocity deviations, respectively. The  $\sigma_p^2$  and  $\sigma_v^2$  represent parameters that can be varied for the study and the values that are used will be stated below. These statistics are used in conjunction with a gaussian random number generator to establish the initial conditions for the actual trajectory.

The noise corrupting the measurements is also assumed to be gaussian and has mean zero and covariance matrix  $R_k$ . The  $R_k$  is treated as having the form

$$R_k = \begin{bmatrix} \sigma_L^2 I & 0 \\ 0 & \sigma_s^2 \end{bmatrix}$$

where the  $\sigma_L^2$  and  $\sigma_s^2$  represent the variance of the noise in the local vertical and subtended angles, respectively. Frequently it will be true that  $\sigma_L^2 = \sigma_s^2$ . The values for these constants will be stated below.

The time interval between measurements provides another parameter for the study. Several different intervals were considered, but only results relating to a sampling interval of 10 minutes will be presented. Since the period of the orbit is approximately 90 minutes, this sampling interval results in measurement data being available at every 40 degrees of subtended arc. The other intervals that were investigated did not measurably change the conclusions

suggested by this sample rate. It should be noted that Meditch [49] has shown that the system is not observable if measurement data are available only at intervals of 180 degrees of subtended arc.

#### 5.2.1 When Linear Theory is Valid

The Kalman filter has been applied [49-53] to the problem of estimating the state of a spacecraft for a variety of missions and has, in general, proved to give satisfactory results. Mendelsohn [54] has discussed a case in which it has not given satisfactory results, however. Many of these studies have dealt entirely with the error covariance matrix  $P_k$  and have not involved any Monte Carlo simulation. Such a procedure is entirely justified if the system were actually linear. But since a nonlinear system is being approximated by a linear system, the validity of  $P_k$  as a measure of the response of the filter depends heavily upon the accuracy of the approximation. In this paragraph, we consider a case in which the linear system (5.6) apparently provides a good approximation to the behavior of the actual trajectory relative to the pre-specified nominal.

For the remainder of this paragraph, let

$$\begin{aligned}\sigma_p^2 &= (5,000 \text{ ft})^2 \\ \sigma_v^2 &= (5 \text{ ft/sec})^2\end{aligned}$$

The initial conditions for the actual trajectory are selected from an ensemble described by these statistics. For the data in Table 5.1, the initial deviations are

$$\underline{x}(t_0) - \underline{x}^*(t_0) = \begin{bmatrix} -2239 \\ 6182 \\ 4192 \\ -3.9 \\ 3.5 \\ -4.0 \end{bmatrix}$$

At any time  $t_k$ , the deviation should be described approximately by

$$\underline{x}_k - \underline{x}_{k-1} \approx \Phi_{k,0}(\underline{x}_0 - \underline{x}_0^*)$$

if (5.6) is adequate. The error in X-component of position and of velocity are depicted in Figure 5.1 for five orbital revolutions for one particular set of initial conditions. This trajectory appears to be representative for the group that were simulated.

It is clear from Figure 5.1 that the error is oscillatory and has an increasing amplitude. In this case the errors do not appear to be significant, so one would expect that linear theory is adequate. To see that this intuitive idea is true, consider the following case.

$$\sigma_L^2 = \sigma_s^2 = (0.1 \text{ degrees})^2$$

Sampling interval = 10 minutes

All of the estimation policies described in Section 5.1.2 were applied to this configuration. Results for Cases (L-1), (L-2), (N-1) and (N-2) are contained in Table 5.1(a). Cases (L-3), (L-4), and (L-5) are described by Table 5.1(b). Only two revolutions are studied for Case (L-5) because of the number of computations that are involved.

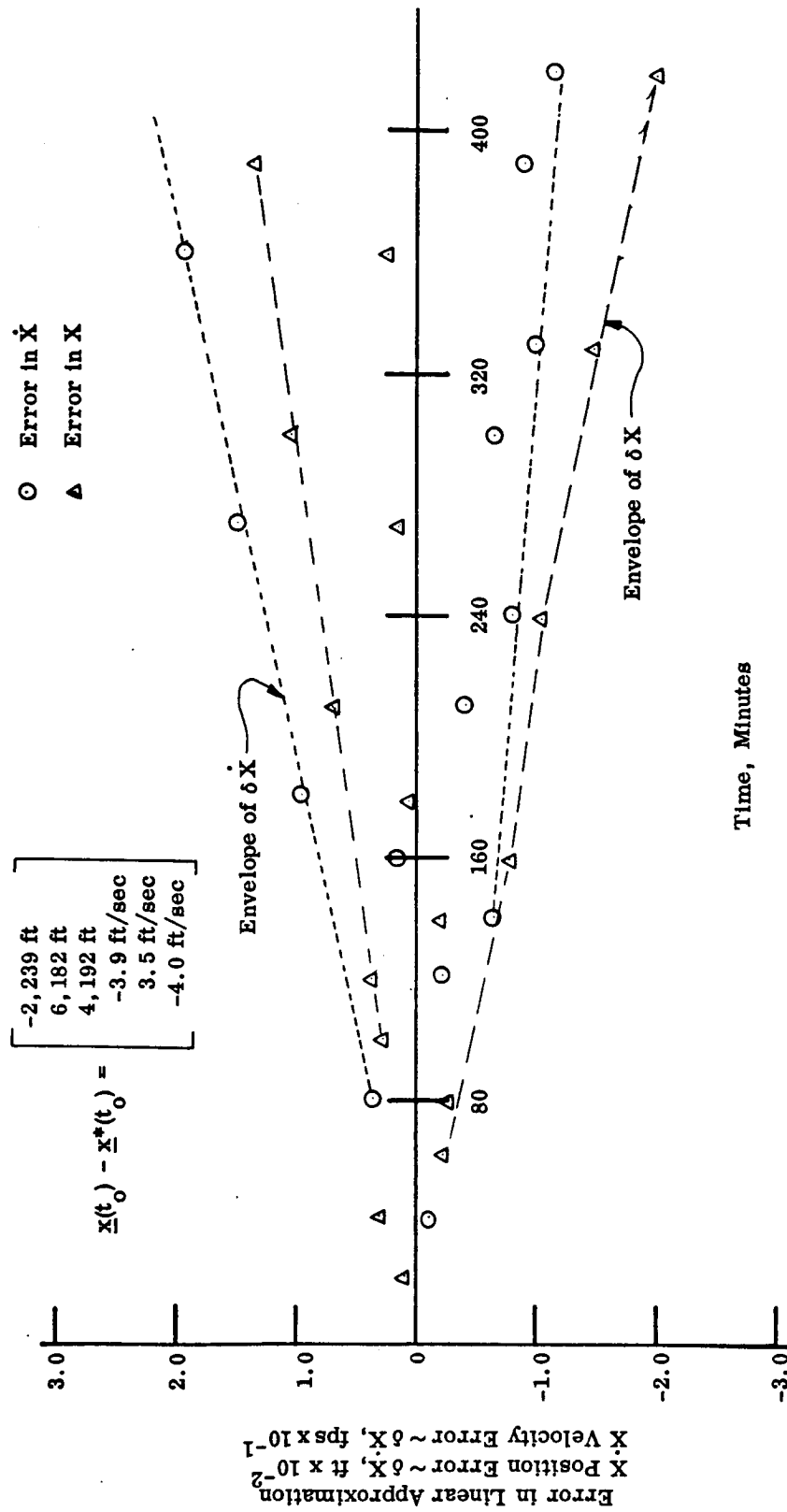


Figure 5.1. Error in Linear Approximation - Case 1

The results in Table 5.1 verify that the linear model is an excellent approximation. First, observe that the axes\* of the error ellipsoids provide an accurate measure of the error in the estimate. Second, it is clear that the results are essentially the same for every filter configuration. As a precursor of things to come, it can be seen that rectification of the nominal does result in a smaller actual error in many cases, particularly for the last two orbital revolutions. The error covariance matrix is unaffected, however.

### 5.2.2 Orbit Rectification to the Rescue

In this paragraph a trajectory is considered for which linear theory proves to be totally inadequate for the description of the state perturbation. For this case, the statistics of the initial perturbation are taken to be

$$\sigma_p = 50,000 \text{ ft.}$$

$$\sigma_v = 50 \text{ ft/sec.}$$

Several different sets of initial perturbations were studied. In the results below, these conditions were

$$\underline{x}(t_0) - \underline{x}^*(t_0) = \begin{bmatrix} 40058 \\ 75488 \\ 35676 \\ -29.8 \\ -125.4 \\ 3.5 \end{bmatrix}$$

In Figure 5.2 the error in the linear approximation in the  $\underline{X}$  and  $\dot{\underline{X}}$  components is depicted. Note that on the scale used, the errors in Figure 5.1 could not be distinguished from the axis of abscissas. From the magnitude of these

---

\* The direction of the axis is not exactly in the direction of the coordinate axes so the magnitudes are listed according to the coordinate axis that is most closely aligned.

TABLE 5.1(a)

Filter Response when the Perturbations are Linear

$$\sigma_p = 5000 \text{ ft}; \sigma_v = 5 \text{ fps}$$

$$\sigma_L = \sigma_s = 0.1 \text{ degree}$$

Time	Component	Prespecified Nominal				Rectification			
		Error in Estimate		Axes of Error Ellipsoids		Error in Estimate		Axes of Error Ellipsoids	
		(L-1)	(N-1)	(L-1)	(N-1)	(L-2)	(L-2)	(L-2)	(N-2)
5400	X, ft	-5639	-5639	2468	2467	-5608	-5605	2478	2478
	Y, ft	16082	16069	16428	16435	16071	16044	16433	16446
	Z, ft	4108	4108	4798	4798	4107	4107	4798	4798
	$\dot{X}$ , fps	-20.7	-20.7	18.7	18.7	-20.7	-20.7	18.7	18.7
	$\dot{Y}$ , fps	0.8	0.8	2.9	2.9	0.8	0.8	2.9	2.9
	$\dot{Z}$ , fps	-6.1	-6.1	4.9	4.9	-6.1	-6.1	4.9	4.9
10800	X, ft	-4573	-4566	2078	2076	-4516	-4505	2100	2098
	Y, ft	3016	2999	14645	14658	2896	2858	14638	14650
	Z, ft	2075	2074	4599	4599	2073	2072	4599	4600
	$\dot{X}$ , fps	-4.3	-4.2	16.6	16.6	-4.1	-4.1	16.6	16.6
	$\dot{Y}$ , fps	3.4	3.4	2.4	2.5	3.5	3.5	2.5	2.5
	$\dot{Z}$ , fps	-7.5	-7.5	4.8	4.8	-7.5	-7.5	4.8	4.8
16200	X, ft	1457	1463	1833	1833	1565	1580	1846	1845
	Y, ft	-9920	-9934	12735	12734	-10056	-10090	12598	12600
	Z, ft	1899	1899	4405	4405	1898	1898	4406	4406
	$\dot{X}$ , fps	10.0	10.0	14.4	14.4	10.2	10.2	14.3	14.3
	$\dot{Y}$ , fps	7.4	7.4	2.2	2.2	7.5	7.5	2.2	2.2
	$\dot{Z}$ , fps	-6.8	-6.8	4.7	4.7	-6.8	-6.8	4.7	4.7
21600	X, ft	-1696	-1695	1661	1660	-1187	-1180	1665	1665
	Y, ft	2258	2251	11322	11316	1625	1618	11214	11211
	Z, ft	497	497	4218	4219	496	496	4219	4220
	$\dot{X}$ , fps	-1.0	-1.0	12.8	12.8	-0.2	-0.2	12.7	12.7
	$\dot{Y}$ , fps	-0.2	-0.2	2.0	2.0	0.3	0.3	2.0	2.0
	$\dot{Z}$ , fps	-7.3	-7.3	4.7	4.7	-7.3	-7.3	4.7	4.7
27000	X, ft	-3442	-3445	1530	1530	-2603	-2559	1535	1535
	Y, ft	3800	3801	10262	10246	3000	2995	10346	10346
	Z, ft	1392	1393	4042	4043	1397	1392	4044	4044
	$\dot{X}$ , fps	-3.3	-3.3	11.6	11.6	-2.3	-2.3	11.7	11.7
	$\dot{Y}$ , fps	-2.0	-2.0	1.8	1.8	-1.1	-1.1	1.8	1.8
	$\dot{Z}$ , fps	-7.7	-7.7	4.7	4.6	-7.7	-7.7	4.7	4.7

TABLE 5.1(b)

Filter Response when the Perturbations are Linear

$$\sigma_p = 5000 \text{ ft}; \sigma_v = 5 \text{ fps}$$

$$\sigma_L = \sigma_s = 0.1 \text{ degree}$$

Time	Component	Prespecified Nominal					Rectification	
		Error in Estimate			Axes of Error Ellipsoid		Error in Estimate	Axes of Error Ellipsoid
			Initial State	Current State				
		(L-3)	(L-4)		(L-3)	(L-4)	(L-5)	(L-5)
5400	X, ft	-5634	-3438	-5623	2468	2468	-5585	2468
	Y, ft	16062	4281	16087	16422	16428	16072	16432
	Z, ft	4109	4734	4110	4798	4798	4108	4799
	$\dot{X}$ , fps	-20.7	1.8	-20.8	18.7	18.7	-20.7	18.7
	$\dot{Y}$ , fps	0.8	3.3	0.7	2.9	2.9	0.7	2.9
	$\dot{Z}$ , fps	-6.1	-5.4	-6.1	4.9	4.9	-6.1	4.9
10800	X, ft	-4576	-3860	-4558	2079	2078	-4485	2090
	Y, ft	3019	5165	3015	14635	14645	2852	14629
	Z, ft	2075	3605	2076	4599	4599	2072	4600
	$\dot{X}$ , fps	-4.3	2.1	-4.3	16.6	16.6	-4.0	16.6
	$\dot{Y}$ , fps	3.4	4.7	3.4	2.5	2.4	3.5	2.5
	$\dot{Z}$ , fps	-7.5	-6.5	-7.5	4.8	4.8	-7.5	4.8
16200	X, ft	1461	-2296	1467	1834	1833	No Data	
	Y, ft	-9943	4628	-9905	12728	12735		
	Z, ft	1899	2917	1900	4405	4405		
	$\dot{X}$ , fps	10.0	1.5	10.0	14.4	14.4		
	$\dot{Y}$ , fps	7.4	3.1	7.4	2.2	2.2		
	$\dot{Z}$ , fps	-6.8	-5.4	-6.8	4.7	4.7		
21600	X, ft	-1679	-1229	-1656	1661	1661	No Data	
	Y, ft	2220	5587	2252	11316	11322		
	Z, ft	498	3474	499	4218	4218		
	$\dot{X}$ , fps	-1.0	3.0	-1.0	12.8	12.8		
	$\dot{Y}$ , fps	-0.1	1.5	-0.2	2.0	2.0		
	$\dot{Z}$ , fps	-7.3	-6.1	-7.3	4.7	4.7		
27000	X, ft	-3354	-1418	-3433	1531	1530	No Data	
	Y, ft	3675	5499	3806	10256	10262		
	Z, ft	1392	3397	1396	4042	4042		
	$\dot{X}$ , fps	-3.2	2.8	-3.3	11.6	11.6		
	$\dot{Y}$ , fps	-1.9	1.7	-2.0	1.8	1.8		
	$\dot{Z}$ , fps	-7.7	-5.1	-7.7	4.6	4.7		



errors, one would suspect that a linear estimation theory could not provide satisfactory results. As shall be shown, this suspicion is entirely valid for the case in which the original nominal is retained throughout the duration of the flight. However, rectification of the nominal provides a striking improvement in the accuracy of the estimates. Once again, consider the following cases

$$\sigma_L^2 = \sigma_s^2 = (0.1 \text{ degrees})^2$$

Sampling interval = 10 minutes

Results for cases (L-1), (L-2), (N-1), (N-2) are stated in Table 5.2(a). It is interesting to examine these data in more detail. First, the error in the estimate for Case (L-1) becomes intolerable and at no time does the error ellipsoid describe the error. This disparity is a manifestation of the nonlinear effects upon the estimated procedure. The error in the estimate during the second and third orbits has a remarkable correlation with the error in the linear approximation. For example, the error in the linear approximation of 16200 seconds is

$$[\underline{x}(16,200) - \underline{x}^*(16,200)] - \Phi(16,200, 0)\delta \underline{x}_0 = \begin{bmatrix} -259,423 \\ -196,697 \\ -2,262 \\ 226.6 \\ -304.1 \\ -6.5 \end{bmatrix}$$

Comparing this with the error in the estimate given in Table 5.2(a), one observes that the error can be attributed almost entirely to the linear approximation error. This is important when considering the effects of the estimation policy (L-4).

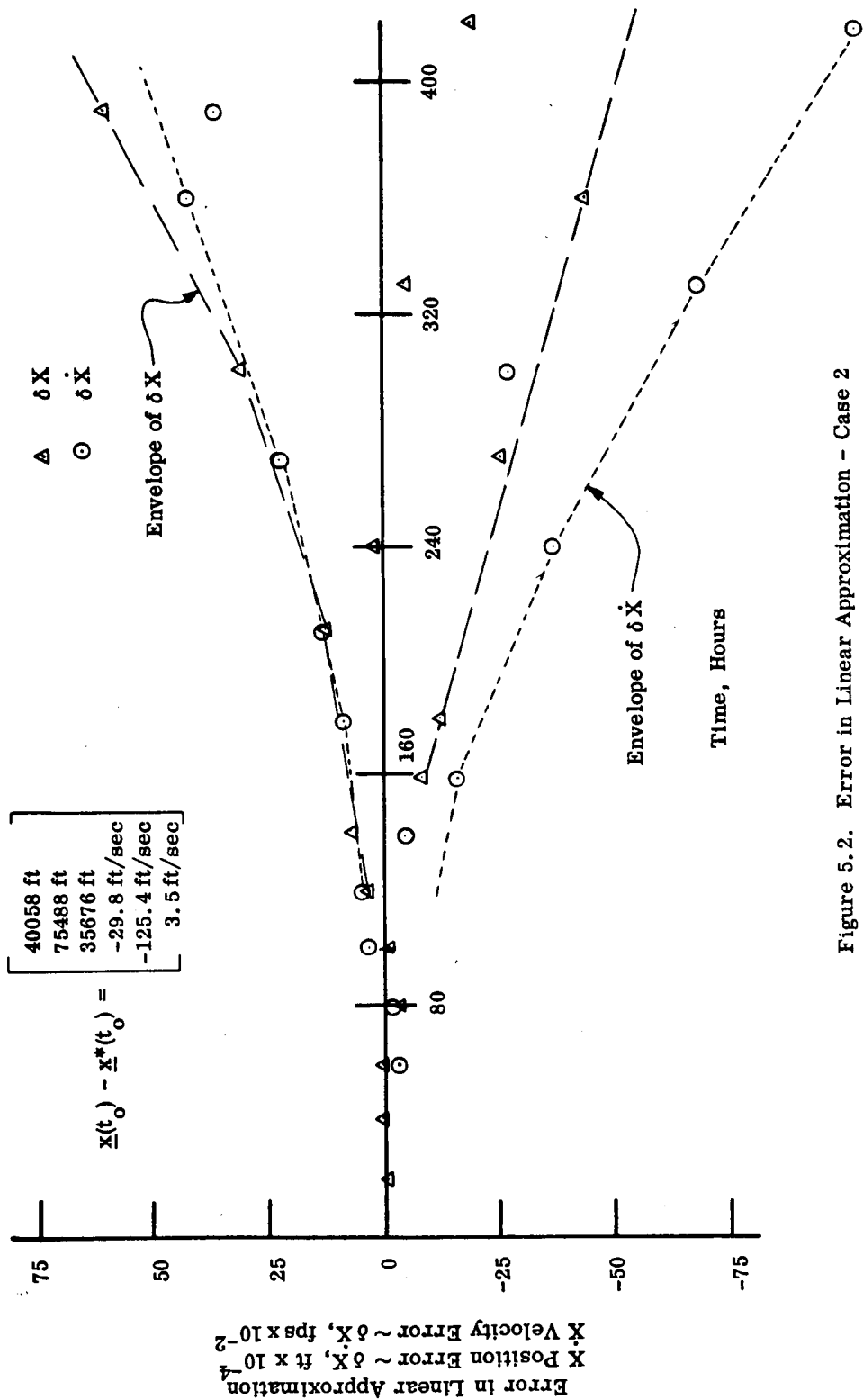


Figure 5.2. Error in Linear Approximation - Case 2

When the nonlinear filter (i. e. , Case (N-1) ) is used, an improvement in the estimate is obtained, particularly during the first orbit. It is apparent that the general behavior of the filter is not greatly affected and that it is the nonlinearities arising in the dynamics that are completely dominant. Again, the error covariance matrix does not provide a valid measure of the actual error in the estimate.

The most significant result appears when examining the effect of rectifying the nominal at each sampling time. For this case, the error in the estimate is greatly reduced compared with Cases (L-1) and (N-1) and does, in fact, correspond with the errors predicted by the error covariance matrix. It is also seen that the error covariance matrix does not appear to be significantly different from the values obtained for the preceding cases.

The introduction of the nonlinear filter has an unexpected effect upon the estimate judging by the tabulated data. In almost every instance, in the table, the error in the estimate is larger although the error covariance matrix is not affected to any great extent. These results suggest a problem that is considered in more detail in Chapter 7. However, the tabulated data are somewhat misleading. For example, an examination of the error in the estimate of the X component of position for the first ten observations indicates that the nonlinear filter provides a more accurate estimate at about half of the observation times (six out of ten). This is typical of the response that is observed for all components of the state vector and leads one to conclude that the nonlinear term does not provide any appreciable benefit.

TABLE 5.2(a)

Filter Response when a Prespecified Nominal is Inadequate

$$\sigma_p = 50,000 \text{ ft}; \sigma_v = 50 \text{ fps}$$

$$\sigma_L = \sigma_s = 0.1 \text{ degree}$$

Time	Component	Prespecified Nominal				Rectification			
		Error in Estimate		Axes of Error Ellipsoids		Error in Estimate		Axes of Error Ellipsoids	
		(L-1)	(N-1)	(L-1)	(N-1)	(L-2)	(N-2)	(L-2)	(N-2)
5400	X, ft	-36783	-34569	4352	4418	6028	7420	4038	4034
	Y, ft	-52744	-38406	18118	18492	-1898	-3176	17306	17284
	Z, ft	10599	9398	16562	15771	10499	10597	16507	16603
	$\dot{X}$ , fps	59.9	50.8	20.1	20.2	3.9	7.0	19.5	19.5
	$\dot{Y}$ , fps	-23.7	-30.9	4.7	4.9	-4.8	-5.2	4.2	4.2
	$\dot{Z}$ , fps	-24.8	-24.2	19.6	18.7	-23.4	-23.5	19.5	19.5
10800	X, ft	-137029	-136329	2852	2941	2078	2742	2528	2528
	Y, ft	-93000	-85475	15504	15550	-3730	-4276	15099	15113
	Z, ft	18163	16819	12034	11425	18521	18615	11991	12013
	$\dot{X}$ , fps	106.9	103.4	17.5	17.4	2.1	3.6	17.2	17.2
	$\dot{Y}$ , fps	-123.2	-127.8	3.3	3.4	0.5	0.6	2.8	2.8
	$\dot{Z}$ , fps	-19.6	-19.2	14.4	13.7	-17.8	-17.8	14.4	14.4
16200	X, ft	-275523	-275099	2271	2356	10845	11244	1989	1987
	Y, ft	-194631	-188665	13620	13620	-14723	-14884	13264	13280
	Z, ft	8883	8553	9915	9410	9490	9517	9873	9890
	$\dot{X}$ , fps	228.5	226.4	15.5	15.2	16.5	17.2	15.1	15.1
	$\dot{Y}$ , fps	-284.2	-288.3	2.6	2.7	8.9	9.1	2.3	2.3
	$\dot{Z}$ , fps	-13.6	-14.1	11.9	11.4	-11.3	-11.4	11.9	11.9
21600	X, ft	-7.x10 <sup>6</sup>	-6.x10 <sup>6</sup>	1946	2061	7128	7348	1700	11983
	Y, ft	11.x10 <sup>6</sup>	10.x10 <sup>8</sup>	12210	11564	-8310	-8281	11961	1699
	Z, ft	4871	5401	8625	8188	6108	6110	8588	8600
	$\dot{X}$ , fps	-12653	-11391	13.9	13.2	8.8	9.3	13.7	13.7
	$\dot{Y}$ , fps	-8068	-7767	2.3	2.4	9.1	9.0	2.0	2.0
	$\dot{Z}$ , fps	-14.5	-14.9	10.4	9.9	-11.7	-11.7	10.4	10.4
27000	X, ft	-11.x10 <sup>6</sup>	-11.x10 <sup>6</sup>	1732	1861	-3567	-3357	11138	11142
	Y, ft	14.x10 <sup>6</sup>	14.x10 <sup>6</sup>	11131	10385	987	1055	1518	1519
	Z, ft	4998	5300	7736	7339	7067	7055	7709	7704
	$\dot{X}$ , fps	-16318	-15027	12.7	11.9	-3.3	-3.0	1.8	1.8
	$\dot{Y}$ , fps	-14123	-13717	2.0	2.2	-1.7	-1.6	12.8	12.8
	$\dot{Z}$ , fps	-7.2	-8.4	9.3	8.9	-4.1	-4.1	9.4	9.4

TABLE 5.2(b)

Filter Response when a Prespecified Nominal is Inadequate

$$\sigma_p = 50,000 \text{ ft}; \sigma_v = 50 \text{ fps}$$

$$\sigma_L = \sigma_s = 0.1 \text{ degree}$$

Time	Component	Prespecified Nominal					Rectification	
		Error in Estimate			Axes of Error Ellipsoid		Error in Estimate	Axes of Error Ellipsoid
		(L-3)	Initial State	Current State	Error Ellipsoid		(L-5)	(L-5)
			(L-4)		(L-4)	(L-4)		
5400	X, ft	126	7280	2618	4233	4352	5189	4058
	Y, ft	-85954	-42255	24762	18534	18118	-1129	17345
	Z, ft	10876	-13085	-9605	16511	16562	10541	16520
	$\dot{X}$ , fps	94.0	51.3	-27.7	20.6	20.1	1.8	19.4
	$\dot{Y}$ , fps	-55.4	-13.0	-18.5	4.5	4.7	-5.7	4.3
	$\dot{Z}$ , fps	-24.8	20.7	14.5	19.5	19.6	-23.5	19.6
10800	X, ft	-63848	16695	11685	2730	2852	1653	2556
	Y, ft	-202314	-49084	9309	16115	15504	-3398	15085
	Z, ft	18615	-21664	-17514	11997	12034	18578	11995
	$\dot{X}$ , fps	222.6	55.1	-9.1	18.2	17.5	1.2	17.1
	$\dot{Y}$ , fps	-139.2	-21.8	-24.0	3.1	3.3	0.5	2.9
	$\dot{Z}$ , fps	-19.6	8.8	17.3	14.4	14.4	-17.8	14.4
16200	X, ft	218704	18586	12733	2159	2271	No Data	
	Y, ft	-235101	-55435	80	14009	13620		
	Z, ft	9876	-12567	-10006	9884	9915		
	$\dot{X}$ , fps	252.1	55.8	-4.3	15.9	15.5		
	$\dot{Y}$ , fps	-319.9	-23.3	-19.6	2.5	2.6		
	$\dot{Z}$ , fps	-13.8	1.6	9.1	11.9	11.9		
21600	X, ft	$-6 \times 10^6$	$4 \times 10^6$	$33 \times 10^6$	1852	1946	No Data	
	Y, ft	$11 \times 10^6$	$36 \times 10^6$	$34 \times 10^6$	11564	12210		
	Z, ft	8294	-10529	-51634	8601	8625		
	$\dot{X}$ , fps	-12135	-4978	$35 \times 10^3$	13.2	13.9		
	$\dot{Y}$ , fps	-7754	-227	$32 \times 10^3$	2.2	2.3		
	$\dot{Z}$ , fps	-15.3	0.4	40.1	10.4	10.4		
27000	X, ft	$-11 \times 10^6$	$5 \times 10^6$	$-37 \times 10^6$	1672	1732	No Data	
	Y, ft	$13 \times 10^6$	$13 \times 10^6$	$-9 \times 10^6$	9746	11132		
	Z, ft	9545	-8284	-54342	7725	7736		
	$\dot{X}$ , fps	-13809	$-5 \times 10^3$	$9 \times 10^3$	11.1	12.7		
	$\dot{Y}$ , fps	-12500	-239	$-38 \times 10^3$	2.0	2.0		
	$\dot{Z}$ , fps	-10.0	-9.5	29.1	9.3	9.3		

The effects of introducing estimation policies (L-3), (L-4) and (L-5) are described in Table 5.2(b). The modification of the observation matrix does not improve the filter behavior according to the data in the tabulation. These data are somewhat misleading because a reduction of the error is actually observed during most of the first orbital revolution. However, the nonlinear plant effects become significant and prevent policy (L-3) from providing a significant improvement.

Estimation policy (L-4) is observed to provide a significant improvement in the estimate of  $\delta \underline{x}_k$  until 21,600 seconds (i.e., through the first four revolutions). It was observed that the accuracy of the estimate deteriorates catastrophically at the end of the fourth revolution. The reason for the sudden deterioration is difficult to explain. The estimate of the initial state is well-behaved until this time, although the error is surprisingly large when compared with the error in the estimate of the current state that is obtained directly from it. It is significant that the error covariance matrix provides a reasonably accurate description of the actual error in the estimate prior to 21,600 seconds. This policy appears to be worthy of further consideration.

Only two revolutions of the trajectory were studied for the smoothing policy (L-5) because of the computational load involved. Use of the smoothed estimate to establish the nominal appears to reduce, generally, the error resulting from policy (L-2). The error covariance matrix is not significantly changed, however, so the improvement would appear to be negligible.

### 5.3 CONCLUSIONS OF THE COMPUTATIONAL STUDY

Several estimation policies have been employed in order to determine the state of a spacecraft moving in nearly circular orbit around the Earth. A variety of trajectories and instrument configurations were examined in conjunction with these policies and the results suggest the conclusions that follow.

- (1) Rectification of the nominal at each sampling time allows a striking and significant extension of the linear theory to cases that suggest the need for more sophisticated filtering techniques when the nominal is restricted to be prespecified.
- (2) When there is no plant noise, continual estimation of the initial state and the subsequent use of this estimate in conjunction with the nonlinear plant equation provides a significant improvement in the estimate of the current state over the linear estimate. With this policy, the nominal is prespecified so the number of on-line computations that must be performed is significantly reduced.
- (3) The nonlinear filter of Chapter 4 provides no useful improvement over the results obtained with a linear filter.

Other conclusions are suggested by the results but not as vividly as the preceding three. Some of these aspects (e.g., Case (L-3) is reconsidered) will be discussed in Chapter 7.

PRECEDING PAGE BLANK NOT FILMED.

## CHAPTER SIX

### APPROXIMATION OF THE A POSTERIORI DENSITY FUNCTION FOR NONLINEAR SYSTEMS

It has been pointed out earlier that knowledge of the a posteriori density  $p(\underline{x}_k/\underline{z}^k)$  provides all of the information required to solve the estimation and control problems. It has also been stated that, in general, it is not possible to determine  $p(\underline{x}_k/\underline{z}^k)$  in a convenient, analytical form from the recurrence relation (IV) that describes the behavior of the density from one sampling time to the next. In this chapter, an approach is presented that provides an approximation of the true density function when  $p(\underline{x}_k/\underline{z}^k)$  is nearly gaussian.

In Section 6.1, the general procedure is discussed and the means by which this procedure is implemented is described. Attention is restricted primarily to the estimation problem but the means of extending this approach to include control terms is described. Relations defining the approximate conditional density are stated in Section 6.2 for a scalar, second order system. Although a scalar system is considered, the relations can be generalized to the multi-variable case without additional conceptual difficulties. The notation required to describe the relations becomes considerably more cumbersome, however. The approximate density function for this system provides insight into the effect of nonlinear terms on the character of the density. These aspects are also discussed.

A means of extending the existing linear theory is discussed in Section 6.3. This discussion is related to estimation policy (L-3) of Chapter 5 and



deals with an unsophisticated means for improving the behavior of the Kalman filter. This technique is exercised in the numerical examples contained in Chapter 7.

## 6.1 THE APPROXIMATION PROCEDURE

The procedure that is proposed here is a generalization of the technique that is commonly used in applying linear estimation and control policies to nonlinear systems and was stated earlier in Section 1.3. Suppose that it is desired to estimate the state of the nonlinear, scalar system

$$x_k = f_k(x_{k-1}) + w_{k-1} \quad (6.1)$$

from measurements described by

$$z_k = h_k(x_k) + v_k \quad (6.2)$$

The  $w_{k-1}$  and  $v_k$  are assumed to be samples from gaussian sequences with known statistics.

To determine the a posteriori density function for this system, the following procedure is suggested.

- (A) Assume that  $f_k$  and  $h_k$  can be written in a Taylor series relative to some nominal values of the state  $x_{k-1}^*$ .

To apply linear theory, one must assume that the perturbations from the nominal can be described by the first order terms of the Taylor series.

- (B) Assume that  $p(x_k/z^k)$  has the same form for every sampling time. When the system is linear, this requirement is satisfied naturally because  $p(x_k/z^k)$  is always gaussian. Densities having this

character have been referred to as being of the "reproducing type" [56].

(C) Assume either

(1)  $f_k$  and  $h_k$  are approximated by a specific number of terms of the Taylor series expansion

or

(2) Allow the number of terms that are to be retained to be determined by the form assumed for  $p(x_k/z^k)$ .

Then, introduce the expansions of  $f_k$  and  $h_k$ , and the density functions for the noise sequences into (IV) and establish recurrence relations for the moments of  $p(x_k/z^k)$  subject to restrictions (1) or (2).

The manner in which (C) is accomplished depends on the form of the density assumed in (B). For example, in Chapter 4,  $p(x_k/z^k)$  was assumed to be gaussian. Then, recurrence relations for the mean  $\hat{x}_k$  and the covariance  $P_k$  were derived under the restriction that only the terms of the Taylor series that permitted the gaussian assumption to be satisfied precisely were to be retained. This procedure led to a generalization of the Kalman filter. In Section 6.2, a gaussian approximation is derived that is different than that of Chapter 4. The differences between the two formulations are seen in Chapter 7 to give significantly different numerical results and to point out the need for caution in the manner in which the approximation is established.

The form that is selected for the density (in accordance with the requirement [B]) is arbitrary. In this discussion, a form is selected that approximates a density function but is not a true density because the approximation can sometimes assume negative values. This form has been chosen because it allows one to make use of the fact that  $p(x_k/z^k)$  should be approximately gaussian for many problems of practical interest.

It is possible to write many density functions as a series of orthogonal polynomials associated with some distribution function [10,37]. When this distribution function is gaussian, the orthogonal polynomials are the Hermite and the resulting expansion is referred to as the Gram-Charlier series. An asymptotic expansion closely related to this is the Edgeworth series [10]. These series will be stated here, but they are discussed in more detail in Appendix C.

Consider a random variable  $\xi$  with a known density function, and let  $x$  be the normalized random variable

$$x = \frac{\xi - m}{\sigma}$$

where  $m$  and  $\sigma$  are the mean and standard deviation of  $\xi$ . Denote the probability density for  $x$  by  $f(x)$  and let  $\psi(x)$  represent the gaussian distribution with mean zero and unit variance. Then, the Gram-Charlier expansion of  $f(x)$  is

$$f(x) = \psi(x) \left[ 1 + \frac{1}{3!} c_3 H_3(x) + \frac{1}{4!} c_4 H_4(x) + \dots \right] \quad (6.3)$$

The  $H_n(x)$  are the Hermite polynomials [10]. They satisfy the recurrence relation

$$H_{n+1}(x) = xH_n(x) - nH_{n-1}(x)$$

where

$$H_0(x) = 1$$

$$H_1(x) = x$$

The coefficients  $c_i$  have been called quasi-moments by Stratonovich [57, 58]

and are defined to be

$$c_3 = -\frac{\mu_3}{\sigma^3}$$

$$c_4 = \frac{\mu_4}{\sigma^4} - 3$$

$$\vdots$$

The  $\mu_k$  are central moments of  $\xi$ .

The Edgeworth expansion is closely related and is given by

$$f(x) = \psi(x) \left[ 1 + \frac{1}{3!} c_3 H_3(x) + \frac{1}{4!} c_4 H_4(x) + \frac{10}{6!} c_3^2 H_6(x) + \dots \right] \quad (6.4)$$

Only terms containing the fourth central moment and less have been included in (6.3) and (6.4). Additional terms are given in Appendix C. Note that the Edgeworth expansion contains one more term than the Gram-Charlier when the series is truncated at this point.

The form of  $p(x_k/z^k)$  required by (B) shall be assumed to be truncations of the Edgeworth expansion. In Section 6.2 the terms stated explicitly in (6.4) are retained. The approximation for fewer terms is obtained immediately from the general result of that section. The truncation of the expansion shall

be considered to result in the higher order moments corresponding with the same order moments of  $\psi(x)$ . This assumption will be used in the determination of the prediction density  $p(x_k/z^{k-1})$ .

The approximation for  $p(x_k/z^k)$  using the procedure (A) - (C) is determined by establishing recurrence relations for the moments appearing in the truncated Edgeworth expansion. The determination of the relations is discussed in Section 6.2.

Attention in this chapter has been restricted to the estimation problem. Thus, in (6.1) the control variables  $u_{k-1}$  do not appear. If the control problem were to be considered, the preceding discussion would remain valid. It would be necessary, however, to assume nominal values for the control and to obtain the Taylor series for  $f_k$  in terms of perturbations in both the state and the control variables. The problem of determining the control law is attacked by the methods of Section 2.1 and is accomplished after the recurrence relations for the moments of  $p(x_k/z^k)$  have been determined. The discussion of Section 4.3 indicates the nature of some of the difficulties that are encountered in trying to establish the control law. Note that the determination of the probability density  $p(x_k/z^k)$  immediately provides the solution of the minimum mean-square estimation problem, whereas knowledge of  $p(x_k/z^k)$  only supplies the information that is required before solution of the control problem can be attempted.

## 6.2 THE CONDITIONAL DENSITY FOR A SECOND ORDER SYSTEM

The procedure described in the preceding section shall be utilized to approximate the a posteriori density for a second order system. Suppose that one must estimate the state of the system with plant

$$x_k = f_k x_{k-1} + g_k x_{k-1}^2 + w_{k-1} \quad (6.5)$$

and measurements

$$z_k = h_k x_k + e_k x_k^2 + v_k \quad (6.6)$$

These equations have an obvious relation to the Taylor series for a general, nonlinear system. Thus it has been assumed that the nominal exists and that (6.5) and (6.6) represent the Taylor series approximation including the second order terms. Thus, we have assumed the form for the plant and measurement equations explicitly. This has been done because it provides a more definitive model than results from the policy in which only the terms that permit the desired form for  $p(x_k/z^k)$  to be achieved are retained.

The derivation of the relations describing the moments involves three basic steps. (This can be reduced to two if the first measurement is not available until  $t_1$ , although the statistics for the initial state are prescribed at  $t_0$ .) In this discussion, a measurement is assumed to be obtained at  $t_0$ , so a relation for  $p(x_0/z_0)$  must be established explicitly as well as general relations for  $p(x_k/z^{k-1})$  and  $p(x_k/z^k)$ . Each of the three steps shall be discussed immediately below and the primary results are stated. A detailed derivation is given in Appendix D.

(1) Determination of  $p(x_o/z_o)$

The initial a posteriori density is obtained from (2.15)

$$p(x_o/z_o) = \frac{p(z_o/x_o)p(x_o)}{p(z_o)} \quad (2.15)$$

The distribution for the initial state has been assumed to be gaussian. The conditional density  $p(z_o/x_o)$  is determined from (6.6) and the distribution for  $v_k$ .

$$\begin{aligned} p(z_o/x_o) &= p(z_o - h_o x_o - e_o x_o^2) \\ &= k_v \exp - \frac{1}{2} \left( \frac{z_o - h_o x_o - e_o x_o^2}{r_o} \right)^2 \end{aligned}$$

Since  $e_o$  is non-zero, this density is obviously not gaussian. This term must be approximated in order to obtain a  $p(x_o/z_o)$  having the form (6.4). This approximation is accomplished in the following manner. The  $p(z_o/x_o)$  can be rewritten as

$$\begin{aligned} p(z_o/x_o) &= k_v \exp - \frac{1}{2} \left( \frac{z_o - h_o x_o}{r_o} \right)^2 \\ &\quad \left\{ \exp - \frac{1}{2} \frac{(-2e_o z_o x_o^2 + 2h_o e_o x_o^3 + e_o^2 x_o^4)}{r_o^2} \right\} \end{aligned} \quad (6.7)$$

For an alternate approach, refer to Chapter 4. Introduce (6.7) and  $p(x_o)$  into (2.15) and rewrite as

$$p(x_o/z_o) = k_{con} \exp - \frac{1}{2} \left( \frac{x_o - \hat{\xi}_o}{\pi_o} \right)^2 \{ \exp B(x_o) \} \quad (6.8)$$

where  $k_{\text{con}}$  represents terms that are independent of  $x_o$  (including  $p(z_o)$ ), and

$$\frac{1}{\pi_o^2} \stackrel{\text{Df}}{=} \frac{h_o^2}{r_o^2} + \frac{1}{m_o^2}$$

$$\hat{\xi}_o \stackrel{\text{Df}}{=} \pi_o^2 \left( \frac{h_o z_o}{r_o^2} + \frac{a}{m_o^2} \right)$$

$$B(x_o) \stackrel{\text{Df}}{=} -\frac{e_o}{2} \left( \frac{-2z_o x_o^2 + 2h_o x_o^3 + e_o x_o^4}{r_o^2} \right)$$

If the measurements were linear, then  $e_o$  would be identically zero and  $\hat{\xi}_o$  and  $\pi_o^2$  would be the mean and variance of  $p(x_o/z_o)$ . Thus, they are a scalar version of the Kalman filter equations.

It is necessary to determine the moments of  $p(x_o/z_o)$  from (6.8) in order to put the a posteriori density into the form required by (6.4). This is not immediately possible because of the last factor. If  $\exp B(x_o)$  can be approximated by a power series, then it becomes a simple matter to determine the moments. One could approximate this term by

$$\exp B(x_o) = 1 + B(x_o) + \frac{1}{2!} B^2(x_o) \quad (6.9)$$

Numerical results indicate that this is not an adequate approximation in many instances. In this chapter, the  $B$  is rewritten in terms of the linear estimate  $\hat{\xi}_o$ . That is, one sees from the definition of  $B$  that it has the form

$$B(x_o) = b_2 x_o^2 + b_3 x_o^3 + b_4 x_o^4$$



It is possible to rewrite this as  $\exp \beta_o \exp B_1(\eta_o)$  where  $B_1(\eta_o)$  is defined to be

$$B_1(\eta_o) = \beta_1 \eta_o + \beta_2 \eta_o^2 + \beta_3 \eta_o^3 + \beta_4 \eta_o^4$$

and

$$\eta_o \stackrel{\text{Df}}{=} (x_o - \hat{\xi}_o)$$

Then, the approximation

$$\exp B = \exp \beta_o [1 + B_1(\eta_o) + B_1^2(\eta_o)] \quad (6.10)$$

is introduced. The exponential  $\exp \beta_o$  does not involve  $\eta_o$ , so it can be included in the constant  $k_{\text{con}}$ . This approximation proves, not unexpectedly, to provide better results than are obtained using (6.9). Using (6.10), (6.8) can be rewritten as

$$p(\eta_o/z_o) = k_{\text{con}} \exp - \frac{1}{2} \left( \frac{\eta_o}{\pi_o} \right)^2 [1 + B_1(\eta_o) + B_1^2(\eta_o)] \quad (6.11)$$

The moments  $E[\eta_o^i/z_o]$  ( $i = 0, 1, 2, 3, 4$ ) are easily determined from (6.11).

The central moments of  $p(x_o/z_o)$  are related to these moments according to

$$E[x_o/z_o] = \hat{\xi}_o + E[\eta_o/z_o]$$

$$\stackrel{\text{Df}}{=} \hat{x}_o$$

$$\begin{aligned} E[(x_o - \hat{x}_o)^2/z_o] &\stackrel{\text{Df}}{=} p_o^2 \\ &= E[\eta_o^2/z_o] - E^2[\eta_o/z_o] \end{aligned}$$

$$\begin{aligned} E[(x_o - \hat{x}_o)^3/z_o] &\stackrel{\text{Df}}{=} \mu_o \\ &= E[\eta_o^3/z_o] - 3E[\eta_o/z_o]p_o^2 - E^3[\eta_o/z_o] \end{aligned}$$

$$\begin{aligned}
E[(x_o - \hat{x}_o)^4/z_o] &\stackrel{\text{Df}}{=} v_o \\
&= E[\eta_o^4/z_o] - 4\mu_o E[\eta_o/z_o] - 6p_o^2 E^2[\eta_o/z_o] - E^4[\eta_o/z_o]
\end{aligned}$$

From (6.4) the approximation to  $p(x_o/z_o)$  is seen to be

$$p(x_o/z_o) = k_o \exp - \frac{1}{2} (\zeta_o)^2 \left[ 1 + \frac{1}{3!} c_3 H_3(\zeta_o) + \frac{1}{4!} c_4 H_4(\zeta_o) + \frac{10}{6!} c_3^2 H_6(\zeta_o) \right]$$

where

$$\zeta_o \stackrel{\text{Df}}{=} \frac{x_o - \hat{x}_o}{p_o}$$

$$c_3 \stackrel{\text{Df}}{=} \frac{\mu_o}{p_o^3}$$

$$c_4 = \frac{v_o}{p_o^4} - 3$$

The complete definition of this density is contained in Table 6.1 and in

Appendix D. To repeat, the principal approximation involved in determining

$p(x_o/z_o)$ , aside from the general form, occurs in the simplification of  $\exp B$ .

This approximation reoccurs in the determination of  $p(x_k/z^k)$  from (IV).

## (2) Determination of $p(x_k/z^{k-1})$

The prediction density is determined from (2.13).

$$p(x_k/z^{k-1}) = \int p(x_{k-1}/z^{k-1}) p(x_k/x_{k-1}) dx_{k-1} \quad (2.13)$$

Assume that  $p(x_{k-1}/z^{k-1})$  has the form prescribed by (6.4). It follows from

(6.5) and the density for  $w_{k-1}$  that

$$p(x_k/x_{k-1}) = k_w \exp - \frac{1}{2} \left( \frac{x_k - f_k x_{k-1} - g_k x_{k-1}^2}{q_{k-1}} \right)^2 \quad (6.12)$$

It becomes clear upon a moments reflection that the integration required by (2.13) cannot be directly accomplished. Fortunately, this is not necessary.

Because the approximation calls for the moments of  $p(x_k/z^{k-1})$ , consider their direct calculation from (2.13). For  $i = 1, 2, 3, 4$ ,

$$E[x_k^i/z^{k-1}] = \int x_k^i p(x_k/z^{k-1}) dx_k \quad (6.13)$$

Substitute (2.13) into (6.13) and iterate the integrals. Then

$$E[x_k^i/z^{k-1}] = \int p(x_{k-1}/z^{k-1}) \left\{ \int x_k^i p(x_k/x_{k-1}) dx_k \right\} dx_{k-1} \quad (6.14)$$

The innermost integration is easily accomplished and produces

$$E[x_k/x_{k-1}] = f_k x_{k-1} + g_k x_{k-1}^2$$

$$E[x_k^2/x_{k-1}] = q_{k-1}^2 + (f_k x_{k-1} + g_k x_{k-1}^2)^2$$

$$E[x_k^3/x_{k-1}] = 3q_{k-1}^2 (f_k x_{k-1} + g_k x_{k-1}^2) + (f_k x_{k-1} + g_k x_{k-1}^2)^3$$

$$E[x_k^4/x_{k-1}] = 3q_{k-1}^4 + 6(f_k x_{k-1} + g_k x_{k-1}^2)^2 q_{k-1}^2 + (f_k x_{k-1} + g_k x_{k-1}^2)^4$$

The moments  $E[x_k^i/z^{k-1}]$  are easily determined from (6.14). To this point, no approximations have been introduced other than the original assumptions associated with the form of  $p(x_{k-1}/z^{k-1})$  and with (6.5) and (6.6). In establishing the  $E[x_k^i/z^{k-1}]$ , it is found that for this problem, the fifth through eighth central moments of  $p(x_{k-1}/z^{k-1})$  are required. These have not been computed (although they could be determined when the first four moments of

the a posteriori density are established). To approximate these moments, we assume that they are identical with the gaussian moments. That is, we let

$$\mu_5 = \mu_7 = 0$$

$$\mu_6 = 5p_k^6$$

$$\mu_8 = 7p_k^8.$$

This constitutes the only additional assumption involved in determining  $p(x_k/z^{k-1})$ . The equations are summarized in Table 6.1 and Appendix D.

### (3) Determination of $p(x_k/z^k)$

The  $p(x_k/z^k)$  is determined from (IV).

$$p(x_k/z^k) = \frac{p(x_k/z^{k-1})p(z_k/x_k)}{p(z_k/z^{k-1})} \quad (\text{IV})$$

The derivation is quite similar to that provided for  $p(x_0/z_0)$  in Section 6.2.1. In that case,  $p(x_0)$  was gaussian, whereas its counterpart in IV (i.e., the  $p(x_k/z^{k-1})$ ) has the Edgeworth form. Thus, the approximation of the factor  $\exp B(x_k)$  that arises through the nonlinearities in the measurement device is performed in exactly the same manner. The algebraic manipulations are more involved in this case because the Hermite polynomials must be rewritten in terms of the centered variable

$$\eta_k \stackrel{\text{Df}}{=} x_k - \hat{\xi}_k$$

where  $\hat{\xi}_k$  is the linear estimate. For the details of the derivation, the reader is referred to Appendix D. The general relations that result are stated in Table 6.1.

The moment relations for simpler density approximations are obtained without difficulty from the tabulated equations. In fact, one obtains another gaussian, but nonlinear, filter directly by eliminating all of the terms that are associated with the Hermite polynomials  $H_3$ ,  $H_4$ , and  $H_6$ . This gaussian filter is distinctly different from that of Chapter 4 and shall be seen to give different (and more satisfactory) results in Chapter 7.

It is hoped that the discussion has been sufficiently clear to indicate the ease with which more terms in the Edgeworth series or in the power series approximation of  $\exp B_1$  could be included. This system illustrates all of the problems that would be encountered in including additional terms in the expansions.

The relations in Table 6.1 that describe the calculation of the central moments of a posteriori densities can be described in general by two terms. The first term represents the moments associated with the linear moments  $\hat{\xi}_k$  and  $\pi_k^2$ . To these moments are added perturbation terms that account for the nonlinearities. It may not be immediately obvious that this is the case for  $p(x_0/z_0)$  and  $p(x_k/z^k)$ , but it is more readily seen for  $p(x_k/z^{k-1})$ . It is particularly interesting to note the effect of plant nonlinearities on the symmetry of the distribution. For the sake of discussion, suppose that  $p(x_{k-1}/z^{k-1})$  is gaussian so that  $\mu_{k-1}$  is identically zero and  $v_{k-1}$  is  $3p_{k-1}^4$ . Then, the  $\mu_{k/k-1}$  has the form

$$\mu_{k/k-1} = 6f_k^2 g_k p_{k-1}^4 + 33f_k^2 g_k \hat{x}_{k-1} p_{k-1}^4 + g_k^3 (24 p_{k-1}^4 \hat{x}_{k-1}^2 - 2 p_{k-1}^6)$$

Naturally, if the plant is linear, then  $g_k$  is identically zero and so

$$\mu_{k/k-1} = 0.$$

However, when  $g_k$  is non-zero, it is clear that  $\mu_{k/k-1}$  in general will be non-zero thereby reflecting the loss of symmetry of the distribution. The symmetry of  $p(x_k/z^k)$  is also destroyed in general by nonlinear measurement terms. The approximation of the a posteriori density by a gaussian distribution (e.g., in the case of the Kalman filter) might be suspected of resulting in a mean value that is biased away from the true mean because of the unsymmetric nature of the true density. This is seen to indeed to the case in some of the numerical results of the next chapter.

It should also be noted that the central moments  $p_k^2$ ,  $\mu_k$ ,  $v_k$  all depend upon the measurement data because the term  $B_1(\eta_k)$  contains the data explicitly in the coefficients. This would have to be taken into account in the derivation of optimal control policies using this density approximation.

Relations defining the moments of the a posteriori density, particularly for  $p(x_k/z^k)$ , can be determined using different approximation techniques. It has already been noted that the  $\exp B(x_k)$  can be written immediately in a power series in  $x_k$  rather than first rewriting it in terms of the centered variable  $\eta_k$ . It is not surprising that the former does not give as good an approximation for an equivalent number of terms in the power series. Another variation is possible in the way in which the  $\hat{\xi}_k$  is defined. One can follow the procedure used in Chapter 4 and include all terms of the second order in  $x_k$  in the  $\hat{\xi}_k$  and  $\pi_k^2$ . Then, one obtains

$$\frac{1}{\pi_k^2} = \frac{h_k^2 - 2e_k z_k}{r_k^2} + \frac{1}{p_{k/k-1}^2}$$

$$\hat{\xi}_k = \pi_k^2 \left( \frac{h_k z_k}{r_k^2} + \frac{\hat{x}_{k/k-1}}{p_{k/k-1}^2} \right)$$

It would appear that this procedure might be the most preferable because the  $B(x_k)$  is reduced to

$$B(x_k) = b_3 x_k^3 + b_4 x_k^4$$

It is no longer necessary to approximate the factor  $\exp b_2 x_k^2$ , and this would seem to be an advantage. Unfortunately, it has been found that this procedure leads to a biased estimate and in many cases, particularly when the variance of the noise is small, provides poorer estimates than are obtained with the Kalman filter. Since the minimum mean-square estimates are theoretically unbiased, the bias must be attributed to the error in the approximation of the density. Thus, one would expect the bias to be reduced as additional terms in the Edgeworth series are included. This has been found to be an accurate description of the behavior that is observed in the numerical studies. Considerations pertaining to errors in the approximation of the density are discussed in Chapter 7.

### 6.3 ON EXTENDING THE USE OF LINEAR FILTERS

In this section we offer a non-rigorous technique for including the effects of terms of greater than first order of the Taylor series representations of the plant and measurement equations. Consider the system (6.1) and (6.2) again.

Assume that a nominal  $x_{k-1}^*$  exists and let (6.1) and (6.2) be represented by

$$\delta x_k = f'_k \delta x_{k-1} + \frac{1}{2} f''_k \delta x_{k-1}^2 + w_{k-1} \quad (6.14)$$

$$\delta z_k = h'_k \delta x_k + \frac{1}{2} h''_k \delta x_k^2 + v_k \quad (6.15)$$

where

$$\delta x_{k-1} \stackrel{\text{Df}}{=} x_{k-1} - x_{k-1}^*$$

$$\delta x_k \stackrel{\text{Df}}{=} x_k - f_k(x_{k-1}^*)$$

$$\delta z_k \stackrel{\text{Df}}{=} z_k - h_k[f_k(x_{k-1}^*)]$$

To apply a linear filter, one must represent the evolution of  $\delta x_k$  by a linear difference equation and  $\delta z_k$  by a linear relation with the state. The obvious procedure for (6.14) and (6.15) is to neglect the second order derivatives  $f''_k$  and  $h''_k$ . Consider an alternative procedure.

At the sampling time  $t_k$ , one has an estimate of the state  $\delta x_{k-1}$ . Let  $\delta x_{k-1}$  be written as

$$\delta x_{k-1} = \delta \hat{x}_{k-1} - \delta \tilde{x}_{k-1}$$

Then

$$\begin{aligned} \delta x_k &= f'_k \delta x_{k-1} + \frac{1}{2} f''_k (\delta \hat{x}_{k-1} - \delta \tilde{x}_{k-1}) \delta x_{k-1} + w_{k-1} \\ &= (f'_k + \frac{1}{2} f''_k \delta \hat{x}_{k-1}) \delta x_{k-1} - \frac{1}{2} f''_k \delta \tilde{x}_{k-1} \delta x_{k-1} + w_{k-1} \end{aligned} \quad (6.16)$$

Now, neglect only the term  $\frac{1}{2} f''_k \delta \tilde{x}_{k-1}$ . As long as  $\delta \tilde{x}_{k-1}$  is small compared with  $\delta \hat{x}_{k-1}$ , one would expect (6.16) to provide a better approximation of the behavior of the plant than when the second order effects are



neglected entirely. Then, the factor  $(f'_k + \frac{1}{2} f''_k \delta \hat{x}_{k-1})$  will serve as the linear plant approximation. The prediction of the state at  $t_k$  is seen to be given by

$$\delta \hat{x}_{k/k-1} = (f'_k + \frac{1}{2} f''_k \delta \hat{x}_{k-1}) \delta \hat{x}_{k-1} \quad (6.17)$$

and the variance is

$$p_{k/k-1}^2 = (f'_k + \frac{1}{2} f''_k \delta \hat{x}_{k-1})^2 p_{k-1}^2 + q_{k-1}^2 \quad (6.18)$$

Compare this result with the more precise relations found in Table 6.1(b).

Suppose that  $p(x_{k-1}/z^{k-1})$  are gaussian so that

$$\mu_{k-1} = 0$$

$$v_{k-1} = 3p_{k-1}^4$$

Then, the predicted mean and variance are seen to be

$$\delta \hat{x}_{k/k-1} = (f'_k + \frac{1}{2} f''_k \delta \hat{x}_{k-1}) \delta \hat{x}_{k-1} + \frac{1}{2} f''_k p_{k-1}^2$$

and

$$p_{k/k-1}^2 = (f'_k + \frac{1}{2} f''_k \delta \hat{x}_{k-1})^2 p_{k-1}^2 + q_{k-1}^2 + \frac{3}{4} f''_k p_{k-1}^4$$

Thus, the approximations (6.17) and (6.18) are seen to cause the terms

$\frac{1}{2} f''_k p_{k-1}^2$  and  $\frac{3}{4} f''_k p_{k-1}^4$  to be neglected from the mean and variance equations,

respectively. In many cases, one would not expect this omission to be significant.

A similar procedure can be employed for  $\delta z_k$ . The predicted estimate  $\delta \hat{x}_{k/k-1}$  is used in rewriting (6.15) as

$$\delta z_k = (h'_k + \frac{1}{2} h''_k \delta \hat{x}_{k/k-1}) \delta x_k + v_k \quad (6.19)$$

The equations (6.16) and (6.19) are linear in the state perturbations. In the use of the linear filter, the plant  $F_k$  and measurement coefficients  $H_k$  can now be treated as

$$F_k \stackrel{\text{Df}}{=} f'_k + \frac{1}{2} f''_k \delta \hat{x}_{k-1}$$

$$H_k \stackrel{\text{Df}}{=} h'_k + \frac{1}{2} h''_k \delta \hat{x}_{k/k-1}$$

Of course, if the nominal is rectified with  $\delta \hat{x}_{k-1}$  at every time  $t_k$ , then after the rectification

$$\delta \hat{x}_{k-1} = \delta \hat{x}'_k = 0.$$

so the second order derivatives are eliminated.

This technique is applied in Chapter 7 and produces some interesting results.

TABLE 6.1

## The A Posteriori Density Function

(a) The A Posteriori Density  $p(x_o/z_o)$  for the Initial Sampling Time

$$p(x_o/z_o) = k_o \exp - \frac{1}{2} (\zeta_o)^2 \left[ 1 + \frac{1}{3!} c_3 H_3(\zeta_o) + \frac{1}{4!} c_4 H_4(\zeta_o) + \frac{10}{6!} c_3^2 H_6(\zeta_o) \right]$$

$$\zeta_o = \frac{x_o - \hat{x}_o}{p_o}$$

$$c_3 = - \frac{\mu_o}{p_o^3}$$

$$c_4 = \frac{\nu_o}{p_o^4} - 3$$

Central Moments:

$$\hat{x}_o = \hat{\xi}_o + E[\eta_o/z_o]$$

$$p_o^2 = E[\eta_o^2/z_o] - E^2[\eta_o/z_o]$$

$$\mu_o = E[\eta_o^3/z_o] - 3E[\eta_o/z_o]p_o^2 - E^3[\eta_o/z_o]$$

$$\nu_o = E[\eta_o^4/z_o] - 4\mu_o E[\eta_o/z_o] - 6p_o^2 E^2[\eta_o/z_o] - E^4[\eta_o/z_o]$$

and

$$E[\eta_o^i/z_o] = \frac{E[\eta_o^i] + E[\eta_o^i B_1(\eta_o)] + E[\eta_o^i B_1^2(\eta_o)]}{1 + E[B_1] + E[B_2]}$$

The expected values are obtained relative to the density

$$p(\eta_o) = \frac{1}{\sqrt{2\pi}\pi_o} \exp - \frac{1}{2} \left( \frac{\eta_o}{\pi_o} \right)^2$$

TABLE 6.1 (continued)

where

$$\frac{1}{\pi_o^2} = \frac{h_o^2}{r_o^2} + \frac{1}{m_o^2}$$

$$\hat{\xi}_o = \pi_o^2 \left( \frac{h_o z_o}{r_o^2} + \frac{a}{m_o^2} \right)$$

(See Appendix D for further details.)

Measurement Nonlinearity Terms:

$$B_1(\eta_o) = \beta_1 \eta_o + \beta_{21} \eta_o^2 + \beta_{31} \eta_o^3 + \beta_{41} \eta_o^4$$

$$\beta_1 = 2b_2 \hat{\xi}_o + 3b_3 \hat{\xi}_o^2 + 4b_4 \hat{\xi}_o^3$$

$$\beta_{21} = b_2 + 3b_3 \hat{\xi}_o + 6b_4 \hat{\xi}_o^2$$

$$\beta_{31} = b_3 + 4b_4 \hat{\xi}_o$$

$$\beta_{41} = b_4$$

where

$$b_2 = \frac{e_o z_o}{r_o^2}$$

$$b_3 = -\frac{h_o e_o}{r_o^2}$$

$$b_4 = -\frac{e_o^2}{2r_o^2}$$

TABLE 6.1 (continued)

(b) The Prediction Density  $p(x_k/z^{k-1})$ 

$$p(x_k/z^{k-1}) = k_{k/k-1} \exp - \frac{1}{2} (\zeta_{k/k-1})^2$$

$$[1 + \frac{1}{3!} c_3 H_3(\zeta_{k/k-1}) + \frac{1}{4!} c_4 H_4(\zeta_{k/k-1}) + \frac{10}{6!} c_3^2 H_6(\zeta_{k/k-1})]$$

$$\zeta_{k/k-1} = \frac{x_k - \hat{x}_{k/k-1}}{p_{k/k-1}}$$

$$c_3 = - \frac{\mu_{k/k-1}}{p_{k/k-1}}$$

$$c_4 = \frac{\nu_{k/k-1}}{p_{k/k-1}} - 3$$

Central Moments:

$$\hat{x}_{k/k-1} = f_k \hat{x}_{k-1} + g_k (p_{k-1}^2 + \hat{x}_{k-1}^2)$$

$$p_{k/k-1}^2 = E[x_k^2/z^{k-1}] - \hat{x}_{k/k-1}^2$$

$$\mu_{k/k-1} = E[x_k^3/z^{k-1}] - 3\hat{x}_{k/k-1} p_{k/k-1}^2 - \hat{x}_{k/k-1}^3$$

$$\nu_{k/k-1} = E[x_k^4/z^{k-1}] - 4\hat{x}_{k/k-1} \mu_{k/k-1} - 6\hat{x}_{k/k-1}^2 p_{k/k-1}^2 - \hat{x}_{k/k-1}^4$$

where

$$\begin{aligned} E[x_k^2/z^{k-1}] &= q_{k-1}^2 + f_k^2 E[x_{k-1}^2/z^{k-1}] + 2f_k g_k E[x_{k-1}^3/z^{k-1}] \\ &\quad + g_k^2 E[x_{k-1}^4/z^{k-1}] \end{aligned}$$

TABLE 6.1 (continued)

$$\begin{aligned}
E[x_k^3/z^{k-1}] &= 3q_{k-1}^2 \hat{x}_{k/k-1} + f_k^3 E[x_{k-1}^3/z^{k-1}] + 3f_k^2 g_k E[x_{k-1}^4/z^{k-1}] \\
&\quad + 3f_k g_k^2 E[x_{k-1}^5/z^{k-1}] + g_k^3 E[x_{k-1}^6/z^{k-1}] \\
E[x_k^4/z^{k-1}] &= 3q_{k-1}^4 + 6q_{k-1}^2 [f_k^2 E(x_{k-1}^2/z^{k-1}) + 2f_k g_k E(x_{k-1}/z^{k-1}) \\
&\quad + g_k^2 E(x_{k-1}^4/z^{k-1})] + f_k^4 E(x_{k-1}^4/z^{k-1}) \\
&\quad + 4f_k^3 g_k E(x_{k-1}^5/z^{k-1}) + 6f_k^2 g_k^2 E(x_{k-1}^6/z^{k-1}) \\
&\quad + 4f_k g_k^3 E(x_{k-1}^7/z^{k-1}) + g_k^4 E(x_{k-1}^8/z^{k-1})
\end{aligned}$$

and

$$\begin{aligned}
E(x_{k-1}^2/z^{k-1}) &= p_{k-1}^2 + \hat{x}_{k-1}^2 \\
E(x_{k-1}^3/z^{k-1}) &= \mu_{k-1} + 3\hat{x}_{k-1} p_{k-1}^2 + \hat{x}_{k-1}^3 \\
E(x_{k-1}^4/z^{k-1}) &= \nu_{k-1} + 4\hat{x}_{k-1} \mu_{k-1} + 6\hat{x}_{k-1}^2 p_{k-1}^2 + \hat{x}_{k-1}^4 \\
E(x_{k-1}^5/z^{k-1}) &= 5\hat{x}_{k-1} \nu_{k-1} + 10\hat{x}_{k-1}^2 \mu_{k-1} + 10\hat{x}_{k-1}^3 p_{k-1}^2 + \hat{x}_{k-1}^5 \\
E(x_{k-1}^6/z^{k-1}) &= 5p_{k-1}^6 + 15\hat{x}_{k-1}^2 \nu_{k-1} + 20\hat{x}_{k-1}^3 \mu_{k-1} \\
&\quad + 15\hat{x}_{k-1}^4 p_{k-1}^2 + \hat{x}_{k-1}^6 \\
E[x_{k-1}^7/z^{k-1}] &= 35\hat{x}_{k-1} p_{k-1}^6 + 35\hat{x}_{k-1}^3 \nu_{k-1} + 35\hat{x}_{k-1}^4 \mu_{k-1} \\
&\quad + 21\hat{x}_{k-1}^5 p_{k-1}^2 + \hat{x}_{k-1}^7 \\
E[x_{k-1}^8/z^{k-1}] &= 7p_{k-1}^8 + 140\hat{x}_{k-1}^2 p_{k-1}^6 + 70\hat{x}_{k-1}^4 \nu_{k-1} + 56\hat{x}_{k-1}^5 \mu_{k-1} \\
&\quad + 28\hat{x}_{k-1}^6 p_{k-1}^2 + \hat{x}_{k-1}^8
\end{aligned}$$

TABLE 6.1 (continued)

(c) The A Posteriori Density  $p(x_k/z^k)$ 

$$p(x_k/z^k) = k_k \exp - \frac{1}{2} (\zeta_k)^2 \left[ 1 + \frac{1}{3!} c_3 H_3(\zeta_k) + \frac{1}{4!} c_4 H_4(\zeta_k) + \frac{10}{6!} c_3^2 H_6(\zeta_k) \right]$$

$$\zeta_k = \frac{x_k - \hat{x}_k}{p_k}$$

$$c_3 = - \frac{\mu_k}{p_k^3}$$

$$c_4 = \frac{\nu_k}{p_k^4} - 3$$

Central Moments:

$$\hat{x}_k = \hat{\xi}_0 + E[\eta_k/z^k]$$

$$p_k^2 = E[\eta_k^2/z^k] - E[\eta_k/z^k]^2$$

$$\mu_k = E[\eta_k^3/z^k] - 3E[\eta_k/z^k]p_k^2 - E^3[\eta_k/z^k]$$

$$\nu_k = E[\eta_k^4/z^k] - 4\mu_k E[\eta_k/z^k] - 6p_k^2 E^2[\eta_k/z^k] - E^4[\eta_k/z^k]$$

$$E[\eta_k^i/z^k] = \frac{E\{\eta_0^i B(\eta_k) P(\eta_k)\}}{E\{B(\eta_k) P(\eta_k)\}}$$

where

$$B(\eta_k) \stackrel{\text{Df}}{=} 1 + B_1(\eta_k) + B_1^2(\eta_k)$$

$$P(\eta_k) \stackrel{\text{Df}}{=} 1 + \frac{1}{3!} c_3 H_3(\eta_k) + \frac{1}{4!} c_4 H_4(\eta_k) + \frac{10}{6!} c_3^2 H_6(\eta_k)$$

TABLE 6.1 (continued)

$$\frac{1}{\pi_k^2} = \frac{h_k^2}{r_k^2} + \frac{1}{p_{k/k-1}^2}$$

$$\hat{\xi}_k = \pi_k^2 \left( \frac{h_k^2 z_k}{r_k^2} + \frac{\hat{x}_{k/k-1}}{p_{k/k-1}^2} \right)$$

The  $B_1(\eta_o)$  are defined in the same manner as in Part I with the trivial change of subscript. The  $H_1(\eta_k)$  and the  $E[\eta_k^i/z^k]$  are defined in Appendix D.



## CHAPTER SEVEN

### NUMERICAL COMPARISON OF LINEAR AND NONLINEAR FILTERING TECHNIQUES

It has already been pointed out that simpler approximations can be obtained from the density approximation described in the preceding chapter by neglecting either terms of the Edgeworth expansion or terms of the power series expansion of  $\exp B(x_k)$ . Each of the approximations in the resulting hierarchy describe a minimum mean-square estimator (i. e., the conditional mean provides the estimate). In this chapter, the adequacy of each density approximation is investigated by examining the behavior of the estimates of the state of a dynamical system. These estimates are compared with those obtained from the linear estimator of Chapter 3, the nonlinear estimator of Chapter 4, and the modified linear estimator of Section 6.3. The system that is considered is simple. Nonetheless, it appears to illustrate the important characteristics as clearly as the more complicated systems that have been investigated.

The mathematical model and the computer program are described briefly in Section 7.1. The parameters and estimation policies that are considered are also discussed. The approximation of the true  $p(x_k/z^k)$  by a gaussian density provides an interesting class of filters. They are discussed in Section 7.2. The filters resulting from the Edgeworth expansion approximation (i. e., the nongaussian version) are investigated in Section 7.3. Conclusions suggested by the results of Sections 7.2 and 7.3 are stated and discussed in Section 7.4.

## 7.1 THE GENERAL PROBLEM

Consider a scalar system with plant described by

$$x_k = x_{k-1} + w_{k-1} \quad (7.1)$$

The state is to be estimated from measurement data that are related to the state according to

$$z_k = x_k^2 + v_k \quad (k = 0, 1, \dots) \quad (7.2)$$

The initial state  $x_0$  and the plant and measurement noise sequences  $w_j$ ,  $v_j$  are gaussian with known statistics.

$$E[x_0] = a$$

$$E[(x_0 - a)^2] = m_0^2$$

$$E[v_k] = 0 = E[w_k] \quad \text{for all } k$$

$$E[v_k^2] = r_k^2 ; \quad E[w_k^2] = q_k^2$$

The variances  $m_0^2$ ,  $r_k^2$ ,  $q_k^2$  will constitute the basic parameters for the study. The  $m_0^2$  and  $r_k^2$  will always be greater than zero, whereas  $q_k^2$  will be set equal to zero in many cases. The actual values assigned to these parameters will be specified later.

The estimation policies are based upon the use of perturbation theory. Two different nominals will be utilized. First, the nominal will be chosen to be  $E[x_0]$ , and this value will be retained throughout the observation policy. As an alternative choice, the nominal will be rectified at each sampling time to be the minimum mean-square estimate (as in Chapter 5). The results obtained with each of these choices for the nominal will be compared. In each case,

the perturbation equations have the form

$$\delta x_k = \delta x_{k-1} + w_{k-1} \quad (7.3)$$

$$\delta z_k = h'_k(\xi_k^*) \delta x_k + \frac{1}{2} h''_k(\xi_k^*) \delta x_k^2 + v_k \quad (7.4)$$

where

$$\delta x_{k-1} = x_{k-1} - x_{k-1}^*$$

$$\xi_k^* = x_{k-1}^*$$

$$\delta x_k = x_k - \xi_k^*$$

A digital computer program was developed to simulate this system and to exercise several different estimation policies. The initial state and the noise sequences were obtained from a gaussian random number generator [61]. All of the computations were accomplished using double precision arithmetic.

The results that appear in subsequent sections represent a single realization of the random sequence. The size of the computer program coupled with the nature of the double precision computations on the particular IBM 7040 that was used precluded the possibility of a complete Monte Carlo simulation. That is, it did not appear to be feasible to obtain the number of runs necessary to compute significant sample means and variances (i.e., apparently at least 1000 realizations are required [31]). Instead, the data presented below represent the behavior of each filter when the same noise sequences are encountered. It is reasonable to expect in the comparison of two filters that the one that gives the better response for a given noise realization will be generally more effective. Enough cases were simulated to indicate

that the results that are presented are representative of the type of behavior that should be expected.

The filter configurations that are examined are to a major extent obtained from the density approximation derived in Chapter 6. As has been mentioned earlier, different density approximations are obtained from the most general expression by eliminating the Hermite polynomial terms (i.e.,  $H_6$ ,  $H_4$ ,  $H_3$ ) and/or by including only one term in the power series expansion of  $\exp B$  rather than two. Additional density approximations are provided by the results of Chapter 4. This gaussian approximation is investigated in Section 7.2 to show the effect of different approaches and to indicate that not all approximations should be expected to provide satisfactory results.

## 7.2 FILTERS BASED UPON A GAUSSIAN DENSITY

Four different filters are investigated in this section; two of them are linear whereas the other two are not.

(G-1) Linear (Kalman filter): This filter is described in Chapter 3.

(G-2) Linear filter with modified system matrices: The Kalman filter is utilized, but the system matrices are modified in accordance with the discussion of Section 6.3. In this example, the linear relation between the measurement and state perturbations is

$$\delta z_k = (h'_k + \frac{1}{2} h''_k \delta \hat{x}'_k) \delta x_k + v_k$$

$$\delta \hat{x}'_k = \delta \hat{x}_{k-1}$$

(G-3) Gaussian nonlinear filter Number 1: The Hermite polynomial terms  $H_3$ ,  $H_4$ , and  $H_6$  are eliminated and the linear moments are modified because of the presence of the nonlinear measurement term  $\exp B$ . See Section 6.2 for further discussion.

(G-4) Gaussian nonlinear filter Number 2: The gaussian approximation of Chapter 4 is considered.

These four filter configurations were exercised with a variety of noise realizations and a priori statistics. The results for a particular noise realization is depicted in Figures 7.1 through 7.3. These data indicate the relative behavior of the filters when acting upon the same measurement data.

The system yielding the results shown in Figure 7.1 contained no plant noise. The standard deviation of the initial perturbation was assumed to be 10 percent of the mean value of the initial state.

$$\begin{aligned} E[x_0] &= 1 \\ E[(x_0 - 1)^2] &\stackrel{\text{Def}}{=} m_0^2 \\ &= 0.01 \end{aligned}$$

The initial perturbation was obtained from a gaussian random number generator and was

$$\delta x_0 = -0.04478927$$

Before an observation is processed, the estimate of the perturbation is zero, so the initial error is  $\delta x_0$ .

In Figure 7.1(a), the measurement noise has a standard deviation equal to 1 percent of the nominal state

$$E[v_k^2] \stackrel{\text{Def}}{=} r_k^2 \\ = 0.0001$$

The standard deviation of the error in the estimate (i.e.,  $p_k$ ) was found to be approximately the same for all configurations. Thus, only the standard deviation from the linear filter (i.e., Case G-1) was plotted.

Note first that the error in the linear estimate exceeds the statistic on the eleventh measurement and remains larger thereafter. Although taken by itself, this behavior is not impossible, it nonetheless describes a common occurrence in the application of linear filters to nonlinear problems. As was noted in Chapter 5, a sufficiently accurate measurement device will often cause the statistic to be a poor measure of the actual error because of the importance of the neglected nonlinear effects [31].

The general character of the response of filters (G-2) and (G-3) is similar to that observed with (G-1), but the magnitude of the error is considerably reduced for these two configurations. It is of further interest to note that (G-2) and (G-3) give essentially the same results.

When the filter of Chapter 4 is used, the results are very disappointing. In fact, the error in the estimate is seen to be considerably increased and suggests that the estimate contains a bias error. Since the conditional mean is theoretically unbiased, a bias error could enter only because the approximation of the true density is poor. One would expect that the error would be decreased by improving the density approximation. Although not discussed at length here, such an improvement was attempted and the response was observed to improve.

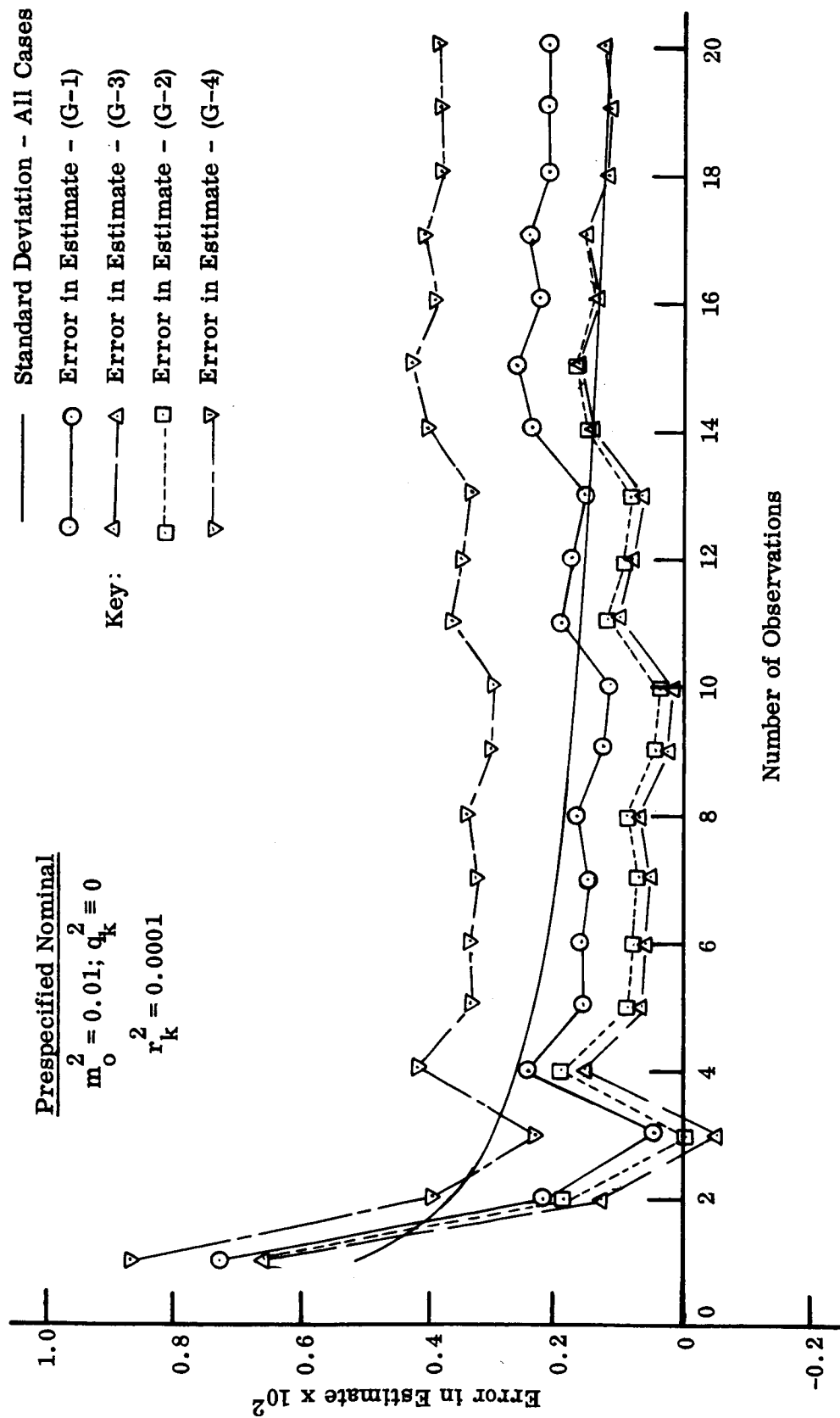


Figure 7.1(a). Comparison of Gaussian Estimators - Noise-free Plant

The results shown in Figure 7.1(a) are based upon a single prespecified nominal. The filter response, when the nominal is rectified at each sampling time, can be seen by referring to Figure 7.1(b). Under this condition, Cases (G-1) and (G-2) are theoretically identical. It is particularly interesting to observe that Case (G-1) in Figure 7.1(b) is essentially identical with Case (G-2) of Figure 7.1(a). Thus, the modification of the system matrices for policies based on a single nominal has the same effect as rectification of the nominal at each sampling time.

Rectification of the nominal is seen to cause all four cases to give basically the same results. This implies that the nonlinear effects are eliminated to a major extent. This becomes more true as the number of samples that have been processed increases.

The standard deviation of measurement noise is increased from 0.01 to 0.1 for the results shown in Figure 7.1(c). One observes that the four filter configurations provide very similar behavior. This result indicates that the magnitude of the measurement noise is more significant than the nonlinearities. Further verification of this statement was provided by finding that rectification does not appreciably influence the results. These data have not been included.

When noise is included in the plant, the filter response is affected in a striking manner. As is known, the presence of plant noise prevents the error variance from vanishing and causes it to approach some nonzero value. As is seen in Figure 7.2, plant noise with a variance of 0.0001 leads to a limiting



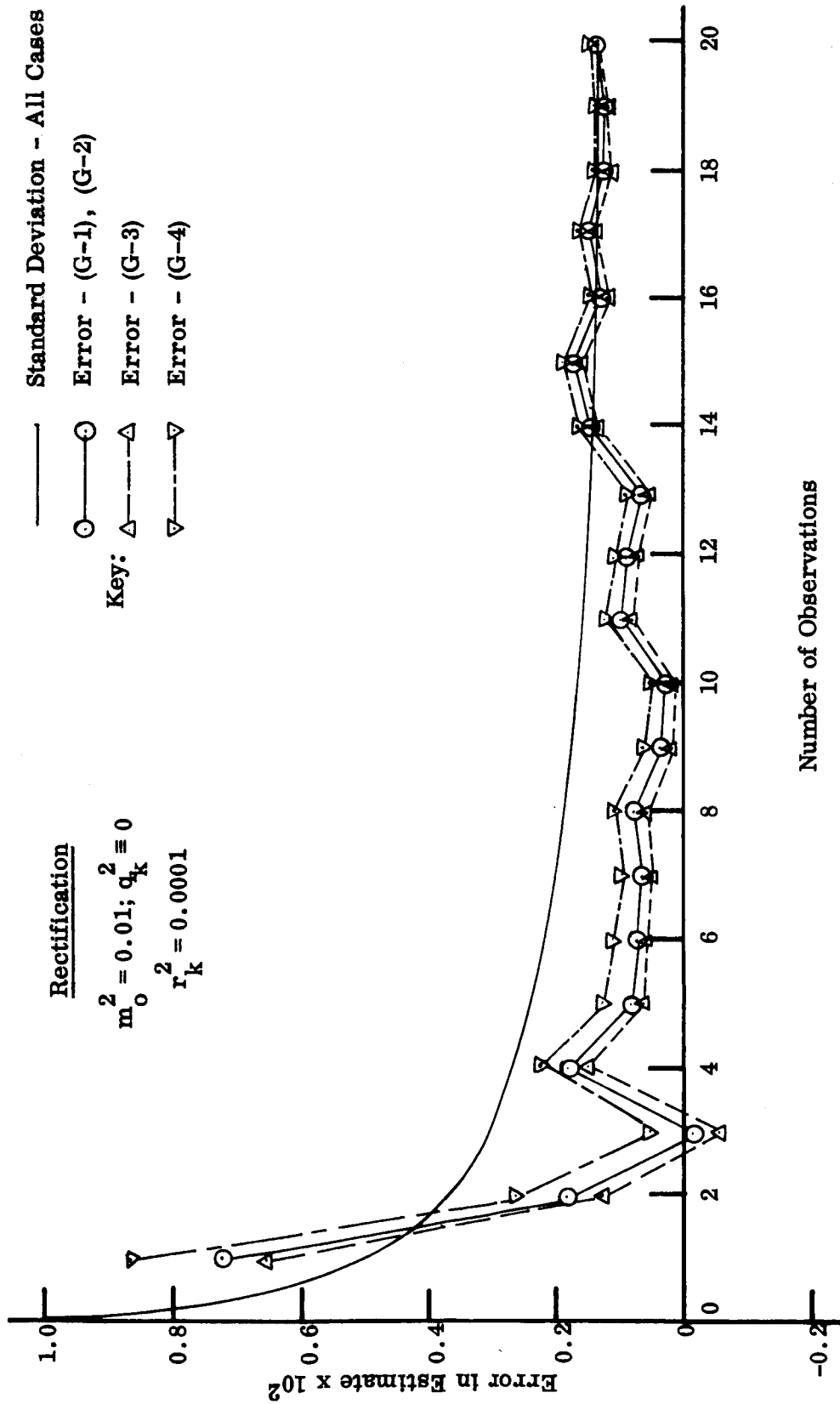


Figure 7.1(b). Comparison of Gaussian Estimators - Noise-free Plant

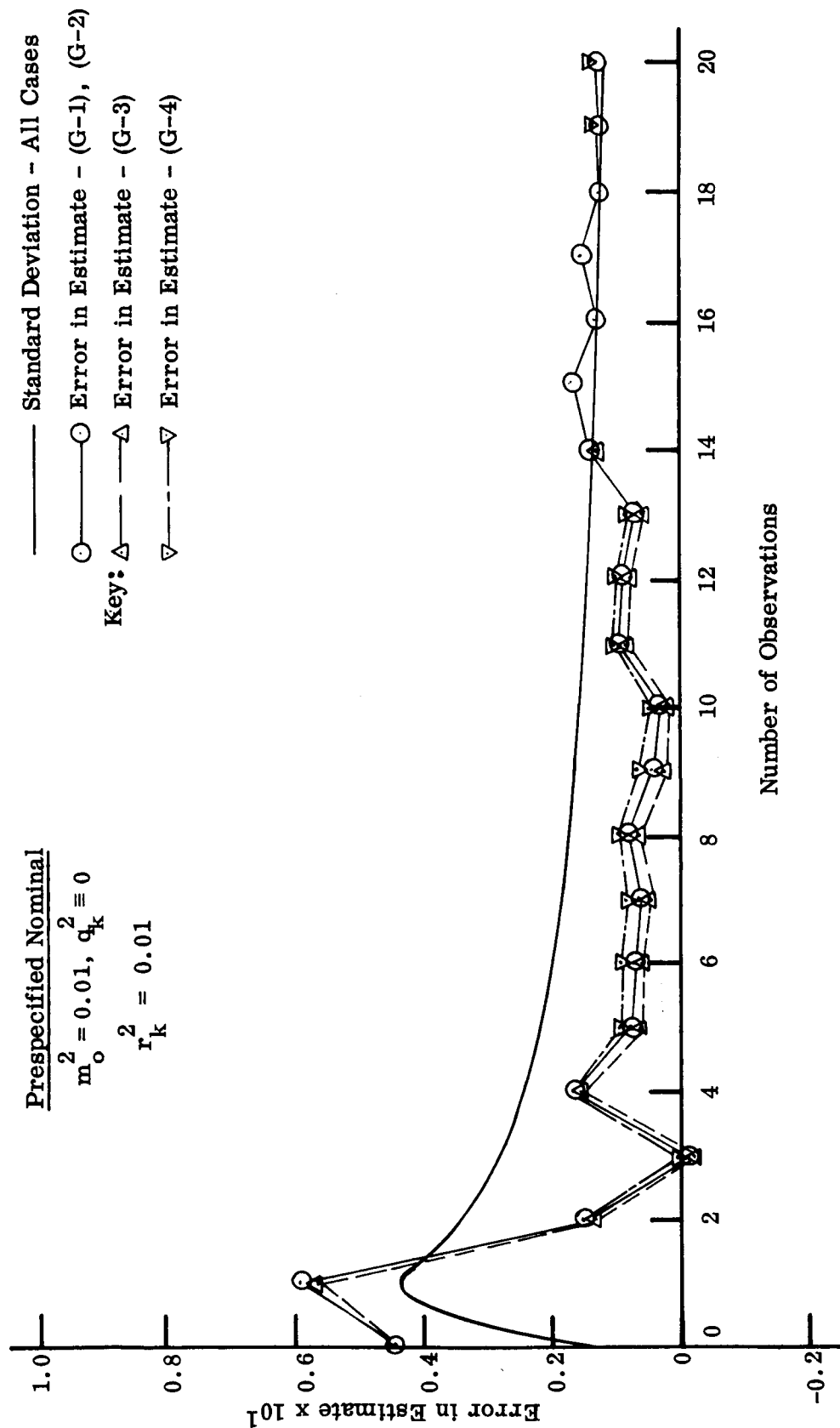


Figure 7.1(c). Comparison of Gaussian Estimators - Noise-free Plant

value of approximately

$$p_{\infty} \approx 0.00455$$

within three samples. One is hard-pressed to judge that one filter provides a more satisfactory response than any other. Naturally, as the level of plant noise decreases, one approaches the behavior shown in Figure 7.1(a). It was also found that rectification does not affect the results of Figure 7.2 in any significant manner.

The nonlinear effects can be amplified by increasing  $m_o^2$  since the initial state perturbation is based upon the initial statistics. To obtain an additional insight into these effects, the  $m_o^2$  was chosen to be

$$m_o^2 = 0.1$$

This led to an initial perturbation of

$$\delta x_o = -0.14163609$$

The behavior observed in Figure 7.1(a) is aggravated by this increased perturbation. The results for a noise-free plant, measurement noise variance of 0.0001, and prespecified nominal are depicted in Figure 7.3. The inadequacy of the standard deviation  $p_k$  as a measure of the error is revealed more clearly than in the preceding data. The linear filter (G-1) and the filter (G-4) are seen to exhibit a definite bias because of the invariance of the error. The fact that (G-4) leads to a deterioration of the accuracy of the estimate is even more apparent. Possibly the most interesting aspect of these data stems from the realization that the modified linear filter (G-2) produces consistently better

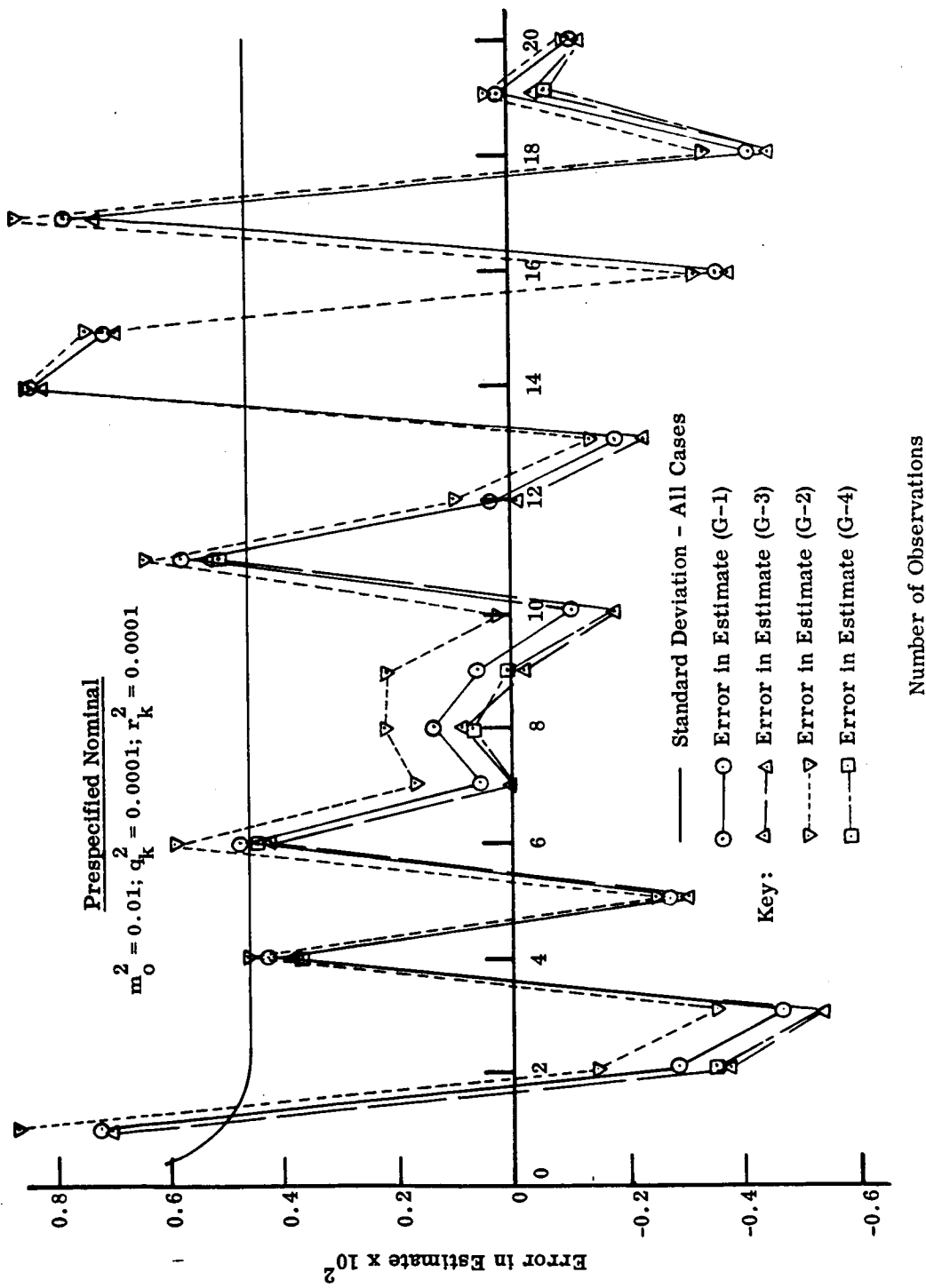


Figure 7.2. Comparison of Gaussian Estimators - Noisy Plant

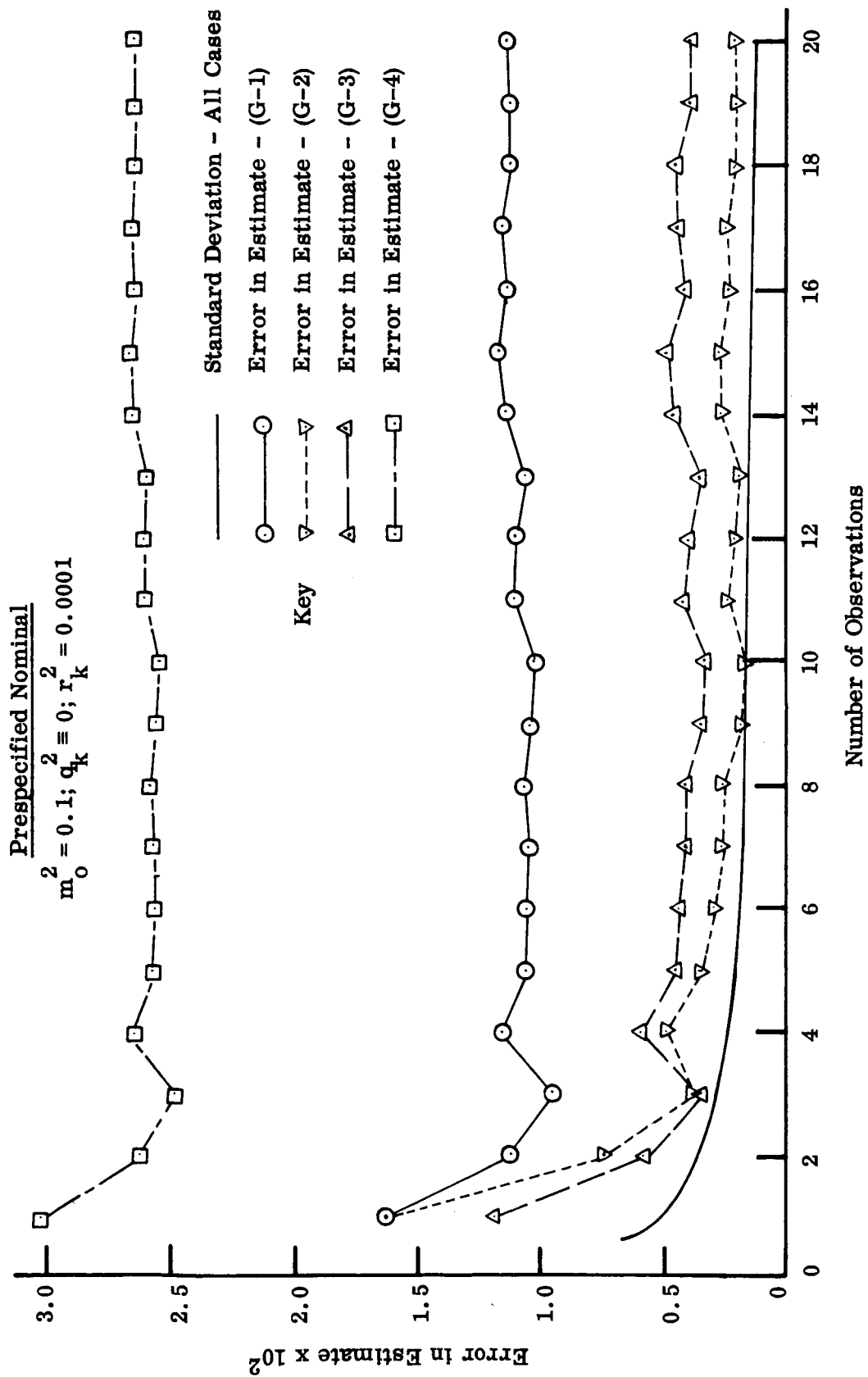


Figure 7.3 Comparison of Gaussian Estimators - Large Initial Perturbation

results than the nonlinear filter (G-3). However, all of the filters lead to results that consistently disagree with  $p_k$ . This case will be discussed further in the next section.

One would expect, judging from past results, that rectification would improve the filter performance. It is noteworthy that when the nominal is rectified, the results do improve and, in fact, the response is described quite adequately by (G-2) in Figure 7.3. Thus, once again for this example, Case (G-2) appears to be equivalent with rectification of the nominal.

### 7.3 ESTIMATORS BASED UPON NONGAUSSIAN DENSITIES

In this section, the Kalman filter is compared with the filters provided by the Edgeworth series approximations. The results are based upon the approximation

$$\exp B_1 \approx 1 + B_1 + \frac{1}{2!} B_1^2$$

where  $B_1$  is defined in Chapter 6. It was found that it was necessary in most instances to use the two-term approximation as opposed to

$$\exp B_1 \approx 1 + B_1$$

in order to obtain sensible behavior for the second and fourth central moments. This aspect is discussed briefly in Section 7.3.2.

The nonlinear filters are identified by the highest order Hermite polynomial that is included. Thus, the estimate provided by most general expansion is referred to as the  $H_6$  filter. Simpler models for the nongaussian densities yield, then, the  $H_3$  and  $H_4$  filters.

### 7.3.1 Nonlinear, Nongaussian Filters

Let us consider the response of this group of nonlinear filters for the cases investigated in the preceding section. In all of these cases, let

$$E[x_0] = 1$$

First, suppose that

$$m_0^2 = 0.01$$

$$r_k^2 = 0.0001$$

$$q_k^2 = 0$$

The initial perturbation is

$$\delta x_0 = -0.04478927$$

and the measurement noise realization is identical with that contained in the data of Figure 7.1.

The results based on the Kalman filter that were included in Figure 7.1 are repeated in Figure 7.4. It was found that the gaussian, nonlinear filter and the  $H_3$  filter provided essentially the same response. This is seen by comparing Figure 7.1 and 7.4. In addition, the  $H_4$  filter exhibited essentially the same response as the  $H_6$  filter. Possibly the most significant difference that is observed is that  $p_k$  for the  $H_4$  and  $H_6$  filters is considerably larger than for any of the lower order filters. This results in a greater sensitivity of the conditional mean to current measurements. This is reflected, unfortunately, by the larger error that is observed during the fourteenth through seventeenth samples. The error is compatible with the statistic  $p_k$ , however.

A prespecified nominal is used in the data for Figure 7.4(a). When the nominal is rectified, the results obtained with all of the filters tend to be essentially the same. These data are depicted in Figure 7.4(b) and again indicate the manner in which rectification seems to eliminate the nonlinear effects.

When the measurement noise is increased (as was done in Figure 7.1(c) ), the linear filter gives the same response as the gaussian, nonlinear filter (G-3) and as the  $H_3$  filter. The  $H_4$  and  $H_6$  filter again are adversely affected during the fourteenth through seventeenth samples.

Plant noise causes the same response for the  $H_3$ ,  $H_4$ , and  $H_6$  filters as was described for the nonlinear gaussian filter (G-3) in the preceding section. These results will not be repeated here.

When the nonlinearity is made more significant by increasing the state perturbation, the nonlinear filters are seen to provide a response that is considerably different from that observed for the gaussian estimators. As seen in Figure 7.5(a), the  $p_k$  for the  $H_3$ ,  $H_4$ , and  $H_6$  filters is larger than for the gaussian filters. Thus, the  $p_k$  appears to be a more accurate measure of the error for these filters. Also note that the  $p_k$  is a random variable for the nonlinear filters, so the  $p_k$  is not as well-behaved. Comparison of Figures 7.3 and 7.5 indicate that the error in the estimate from the modified linear filter (G-2) is subject to less violent changes, but that the  $H_i$  filters give estimates that appear to cope with the nonlinearity more adequately. These filters are more sensitive to the actual measurement noise, however.



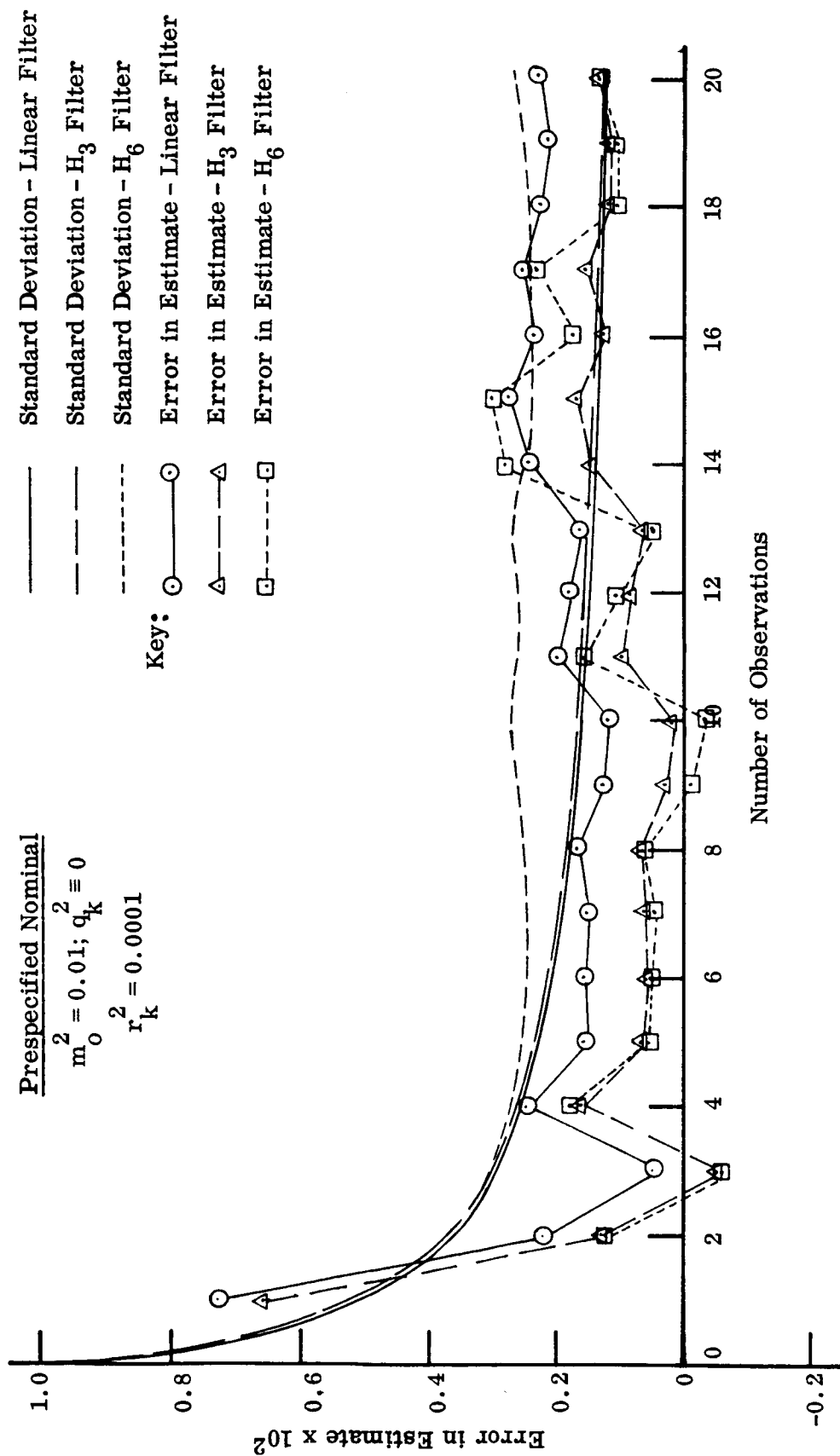


Figure 7.4(a). Comparison of Gaussian and Nongaussian Estimators - Noise-free Plant

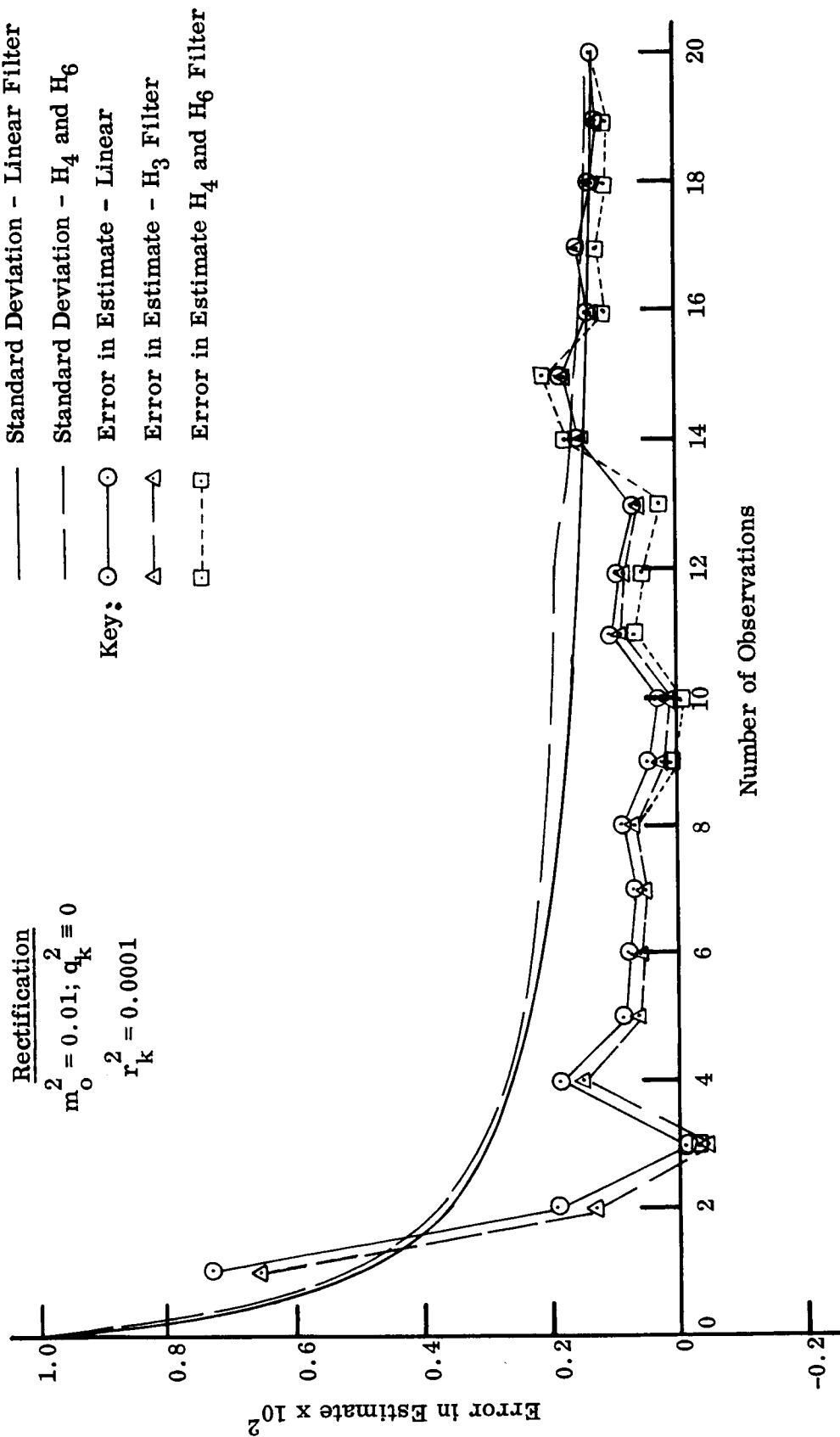


Figure 7.4(b). Comparison of Gaussian and Nongaussian Estimators - Noise-free Plant

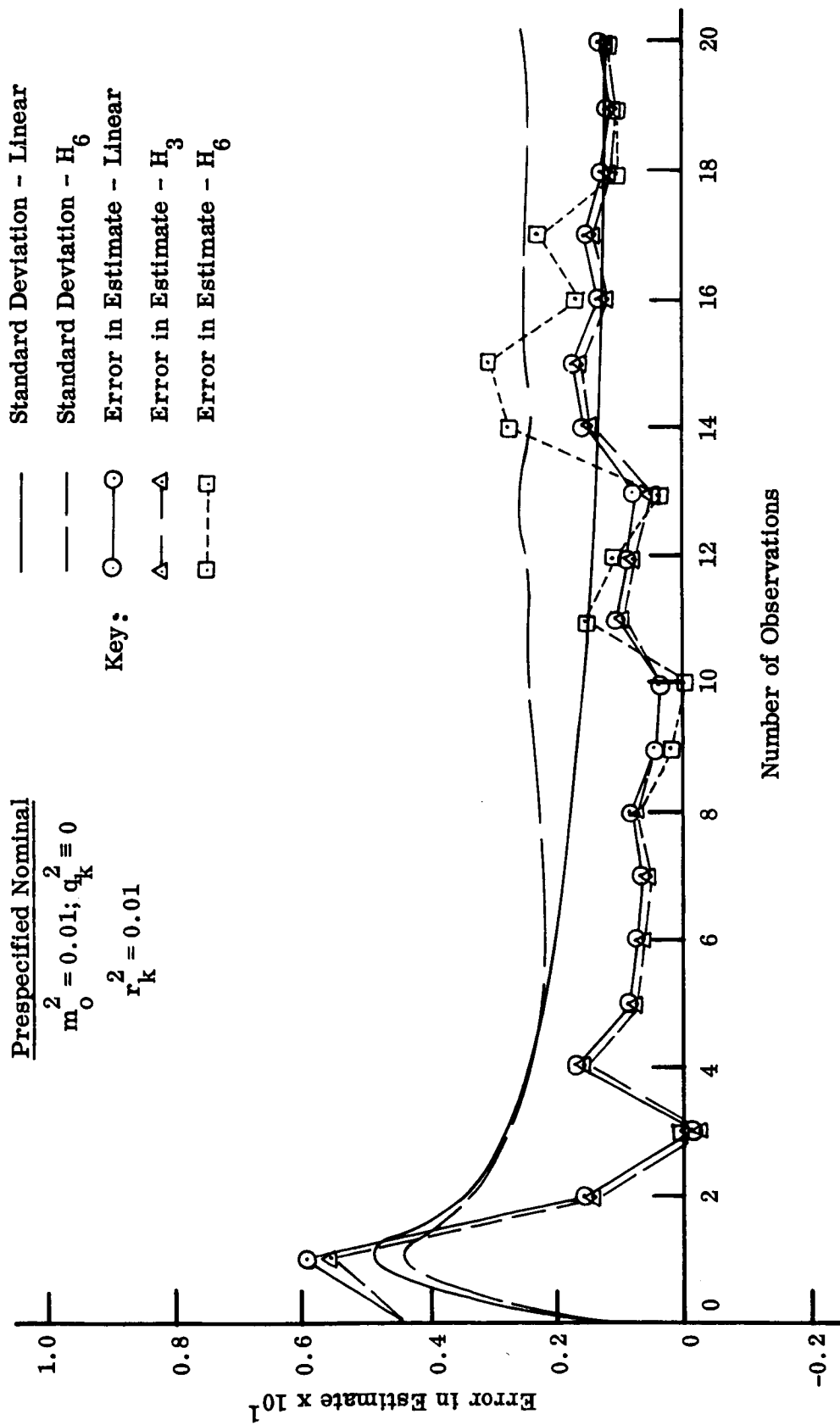


Figure 7.4(c). Comparison of Gaussian and Nongaussian Estimators - Noise-free Plant

Orbit rectification improves the estimate for every case but particularly for the linear filter. The  $H_4$  and  $H_6$  filters appear to provide the better response during a major portion of the interval. However, these filters again are more sensitive to the large measurement noise values that are present in the fourteenth through seventeenth samples. These results are presented in Figure 7.5(b).

### 7.3.2 A Poor Density Approximation

One difficulty manifested itself in some of the cases that were simulated although not in any of the results that have been presented above. It was found that the second and fourth central moments occasionally assumed negative values. This was observed for every one of the nongaussian filters, although it occurred most frequently among the approximations that used

$$\exp B_1 = 1 + B_1.$$

The inclusion of the second order term eliminated the problem to a major extent, although it did not eliminate it entirely when a large state perturbation was experienced. This suggests that it might be appropriate to include the third order term (i.e.,  $1/3! B_1^3$ ) in order to improve the approximation of  $\exp B_1$ . Alternatively, the negative moments could be interpreted as implying that additional terms of the Edgeworth expansion should be included in order to improve the approximation of  $p(x_k/z^k)$ . As has been pointed out, the truncated Edgeworth expansion is not a true density because negative values are assumed for some values of  $x_k$ . This could result in erroneous values for the moments.

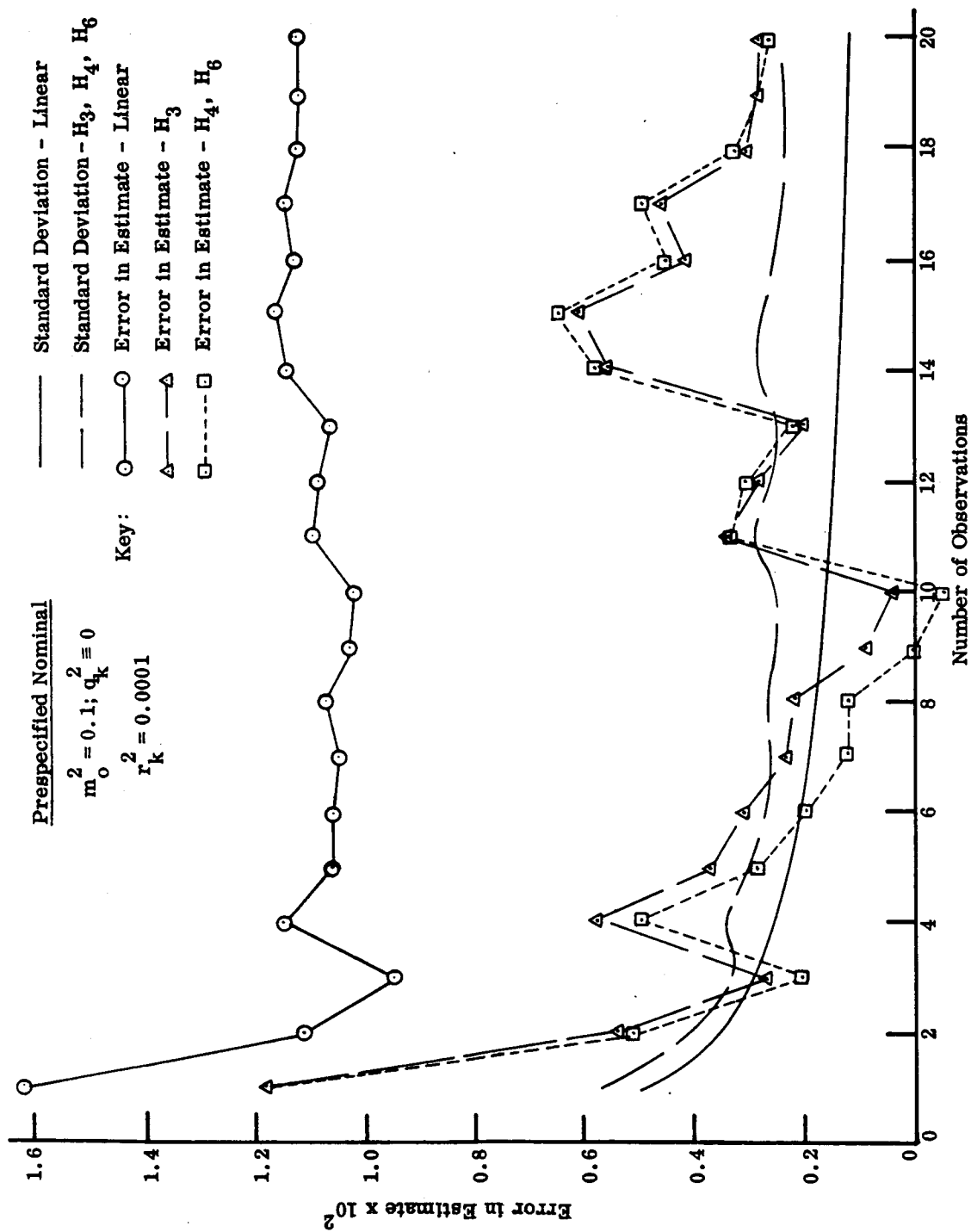


Figure 7.5(a). Filter Response for Large Initial Perturbations

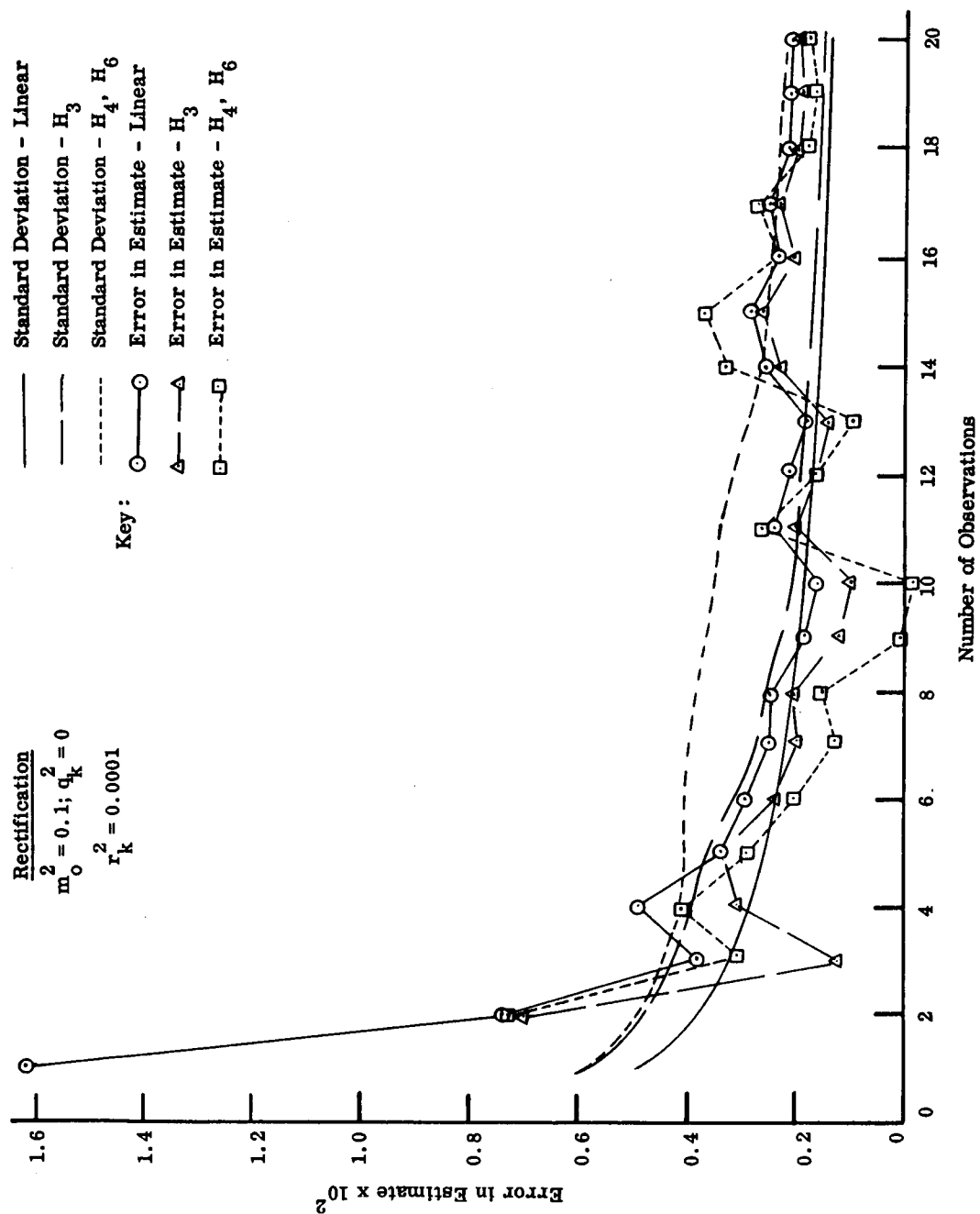


Figure 7.5(b). Filter Response for Large Initial Perturbations

#### 7.4 SUMMARY OF RESULTS

Several conclusions can be drawn from the numerical results of the preceding sections. Although the data are based on a very simple system, more complicated systems have been investigated and appear to corroborate the conclusions that are stated below.

Before proceeding to any statements regarding the merits and demerits of the various estimation policies, it is necessary to recognize the following.

- (0) Unless the measurement noise and/or the plant noise is "small", the linear filter gives essentially the same result as the considerably more complicated nonlinear filters. No attempt shall be made to clarify the circumstances which one can determine if the noise is sufficiently "small" to warrant consideration of the nonlinear filter.

With this provision in mind, it is possible to consider the relative behavior of the filter configurations.

Of the three conclusions stated at the end of Chapter 5, the one dealing with rectification is further substantiated in this chapter, and the one dealing with the nonlinear filter of Chapter 4 can be strengthened.

- (1) Rectification of the nominal at each sampling time causes the behavior of the linear filter to be considerably improved. After a sufficiently large number of measurement samples, the linear filter yields essentially the same response as considerably more complicated systems.

- (2) The gaussian, nonlinear filter of Chapter 4 is not satisfactory.

It does, in fact, appear to yield an estimate that contains a bias which causes it to yield poorer results than even the linear filter.

Thus, one must be cautious when deriving an approximation of a posteriori density.

Several other significant results were obtained. The filter configuration (G-2) based upon modification of the linear system matrices and concomitant application of the linear filter provided a number of suggestive results.

- (3) The filter configuration (G-2) yields the same behavior for the prespecified nominal that is observed when rectification of the nominal is introduced. This estimation policy was utilized in Chapter 5, but did not provide such striking results. This can be attributed to the fact that the plant nonlinearities were not included in the modification of the system matrices because of the inherent difficulties. As was observed, it was these effects that were dominant, however.

- (4) The gaussian nonlinear filter obtained from the general approximation of Chapter 6 yielded results that are comparable with (G-3). Since the filter (G-3) involves many more computations, one would question its usefulness. It might prove useful in the determination of a control policy if one were dealing with that problem.

The Edgeworth approximation provided three filter configurations.

Their behavior shall be summarized in the following manner.



- (5) The  $H_4$  and  $H_6$  filters yielded essentially identical results, whereas  $H_3$  can be more closely identified with the gaussian nonlinear filter (G-3).
- (6) The conditional variance is generally larger than that for the linear and other lower order filters, and in many cases appears to be a more adequate measure of the error in the estimate.
- (7) The estimation error for the  $H_6$  filter is more sensitive to the measurement noise realization than is the lower order filters. This would appear to be a manifestation of the larger values observed for the variance.

The sensitivity of the filter response to the method of approximation is certainly an important consideration. For the example discussed in this chapter, the following conclusions become apparent.

- (8) The approximation of the nonlinear measurement effects (i.e., the factor  $\exp B_1$ ) is of supreme importance. It was found that the  $B_1$  should be written in terms of the centered variable

$$x_k - \xi_k$$

where  $\xi_k$  is the linear estimate rather than leaving it as a polynomial in  $x_k$ . Furthermore, the number of terms retained in the power series approximation of  $\exp B_1$  also has a significant effect. In general, at least terms including the quadratic must be included.

It was found that the quality of the approximation of  $\exp B_1$  could be judged to an extent by the behavior of the second and fourth central moments  $\mu_k$  and  $\nu_k$ . These quantities, which should be positive, sometimes assumed negative values. This was particularly true when  $\exp B_1$  was approximated by the first order terms of the power series.

## CHAPTER EIGHT

### SUMMARY AND CONCLUDING REMARKS

The general problem of determining the optimal control policy for a stochastic, time-discrete, dynamical system was posed in Chapter 1. The general solution of this problem, assuming knowledge of the a posteriori density  $p(\underline{x}_k/\underline{z}^k)$ , was presented in Chapter 2, but more specialized problems were considered thereafter. Attention was restricted primarily to the problem of approximating the density  $p(\underline{x}_k/\underline{z}^k)$ . The adequacy of each approximation was evaluated by examining the behavior of the resulting minimum mean-square estimate. Estimates from both linear and nonlinear filters were compared. The class of linear filters that were considered included several that are based on techniques for extending the range of applicability of the general linear theory.

In the following section, the problem dealt with in each of the preceding chapters is described, and the principal results are summarized. The major conclusions of the study and suggestions for future research are discussed in Section 8.2.

#### 8.1 SUMMARY

CHAPTER 1: The optimal stochastic control problem is described in the following manner. Suppose that the state of a dynamical system evolves according to

$$\underline{x}_k = \underline{f}_k(\underline{x}_{k-1}, \underline{u}_{k-1}, \underline{w}_{k-1}) \quad (I)$$

where  $\underline{u}_{k-1}$  is the control vector and  $\underline{w}_{k-1}$  is a random sequence with known

statistics. The control policy is to be based upon measurement data described by

$$\underline{z}_k = \underline{h}_k(\underline{x}_k, \underline{v}_k) \quad (\text{II})$$

where  $\underline{v}_k$  is another random sequence, independent of the  $\underline{w}_{k-1}$ , with known statistics. Then, choose the control vectors to minimize the expected value of

$$V_N = \sum_{i=1}^N W_i(\underline{x}_i, \underline{u}_{i-1}) \quad (\text{III}_E)$$

The general approach to the solution of this problem and the related problem of state estimation is then described.

CHAPTER 2: The general stochastic control problem stated in Chapter 1 is considered. In particular, it is shown that knowledge of the a posteriori density  $p(\underline{x}_k/\underline{z}^k)$  provides the general solution of the minimum mean-square error estimation problem and the stochastic control problem. The  $p(\underline{x}_k/\underline{z}^k)$  is then shown to evolve according to an integral recurrence relation. Relations describing the prediction and smoothing densities  $p(\underline{x}_{k+\gamma}/\underline{z}^k)$  and  $p(\underline{x}_{k-\gamma}/\underline{z}^k)$  are also derived. In some instances, the application of the general relations to a specific system is simplified by working with the equivalent relations involving characteristic functions. These relations are derived.

CHAPTER 3: The linear, gaussian, optimal stochastic control problem is considered. It is shown that the solution of this problem separates into the dissimilar problems of state estimation and deterministic, optimal control. Furthermore, it is seen that the separation occurs because the error covariance matrix of the estimation problem does not involve the measurement data.

The solutions of the three aspects of the linear estimation problems (i. e. , smoothing, filtering, and prediction) are presented.

CHAPTER 4: Attention is restricted primarily to the problem of approximating the a posteriori density for a system with the plant equation

$$\underline{x}_k = \underline{f}_k(\underline{x}_{k-1}) + \underline{w}_{k-1} \quad (\text{I-N})$$

and the measurement data

$$\underline{z}_k = \underline{h}_k(\underline{x}_k) + \underline{v}_k \quad (\text{II-N})$$

The noise sequences are assumed to be gaussian and the approximation is established under the constraint that the density  $p(\underline{x}_k/\underline{z}^k)$  is gaussian. This procedure leads to a generalization of the Kalman filter of Chapter 3. It is seen that the error covariance matrix becomes a function of the measurement data. Furthermore, it is seen that the choice of the most recent minimum mean-square estimate as the nominal leads to a considerable simplification of the equations.

The problem of determining the control policy for a linear plant from measurements that bear a nonlinear relation to the state is considered briefly. It is suggested that the Separation Principle for completely linear systems is no longer valid because of the dependence of the error covariance matrix upon the measurement data.

CHAPTER 5: The problem of estimating the state of a spacecraft in a nearly circular orbit from the angular measurements provided by a horizon sensor is considered. A nominal is assumed and the results of Chapters 3 and 4 are

utilized in this numerical investigation. Several estimation policies are implemented. Briefly, these policies are:

- (1) Kalman filter using a single prespecified nominal
- (2) Kalman filter using the  $\hat{\underline{x}}_{k-1}$  as the nominal at each sampling time  $t_k$  (the policy of updating the nominal is referred to as rectification)
- (3) Kalman filter with modified system matrices and prespecified nominal (the manner in which the system matrices are modified is discussed in more detail in Section 6.3)
- (4) Continual re-estimation of the initial state using linear theory with subsequent updating of the estimate of  $\underline{x}_k$  using the nonlinear plant equations
- (5) Rectification using a smoothed estimate
- (6) Nonlinear filter of Chapter 4 (referred to as (G-4) with a pre-specified nominal)
- (7) Nonlinear filter (G-4) with rectification.

Of these seven estimation policies, rectification was found to significantly extend the range of linear theory and provided the most satisfactory results in those cases in which policy (1) was found to be inadequate. Policy (4) was also seen to provide excellent results and has the advantage over the rectification policy that the system matrices do not have to be recomputed at each sampling time. It was established that the estimates provided by the nonlinear filter did not, in general, differ from those obtained with the linear filter. Only the

nonlinear measurement effects were included, however, so this apparent ineffectualness could be attributed in part to the dominant role played by the plant nonlinearities. This filter is discussed further in Chapter 7.

CHAPTER 6: A procedure for approximating the a posteriori density  $p(x_k/z^k)$  is proposed. It is suggested that this procedure be implemented using a truncation of an Edgeworth series to approximate the density at each sampling time. The  $p(x_k/z^k)$  associated with a second order, scalar system

$$x_k = f_k x_{k-1} + g_k x_{k-1}^2 + w_{k-1}$$

$$z_k = h_k x_k + e_k x_k^2 + v_k$$

where  $w_{k-1}$  and  $v_k$  are gaussian is approximated by establishing recurrence relations for the first four moments of the distribution. It is shown that the moments for the prediction density  $p(x_k/z^{k-1})$  can be obtained without difficulty. It is necessary to approximate the nonlinear measurement effects which appear as a factor  $\exp B$ . This is accomplished by expressing  $B$  as a polynomial in terms of the variable

$$\eta_k = x_k - \hat{\xi}_k$$

where  $\hat{\xi}_k$  is the linear estimate of  $x_k$ , and then expressing  $\exp B_1$  as a power series

$$\exp B_1 = 1 + B_1 + \frac{1}{2!} B_1^2 + \dots$$

Several different approximations can be obtained from the general result of this chapter including a gaussian approximation that differs from that found in Chapter 4.

Finally, a technique is suggested for improving the approximation provided by linear theory. This procedure results in the use of second order terms of the Taylor series expansion of the plant and measurement equations.

CHAPTER 7: The filters deriving from the density approximations of Chapters 3, 4, and 6 are applied to the problem of estimating the state of the system.

$$x_k = x_{k-1} + w_{k-1}$$

$$z_k = x_k^2 + v_k$$

It was found that the variance of the plant and/or the measurement noise had to be "sufficiently" small before the nonlinear filters provided results that differed significantly from those obtained with the linear filter. In the cases in which the linear and nonlinear filters gave different results, the response of the linear filter was benefitted significantly by rectification of the nominal. Furthermore, the filter of Chapter 4 was observed to yield generally unsatisfactory results. The estimate provided by this approximation appeared to be biased.

The gaussian filter obtained from the general results in Chapter 6 gave satisfactory results in most cases but did not perform significantly better than the modified linear filter. In fact, the modified linear filter appeared to provide the same behavior with a single prespecified nominal that the linear filter exhibited with rectification. The two policies result in essentially equivalent response.



The most general nonlinear filter consistently yielded values for the conditional variance that were larger than those obtained from the linear filter. In many cases, the larger values appeared to more adequately describe the error. This filter was more sensitive to the magnitude of the actual measurement noise realization. A more complete summary and discussion of the numerical results is given in Section 7.4.

## 8.2 GENERAL CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

A procedure for approximating the a posteriori density function has been suggested and has been seen to lead to a straightforward means of accomplishing this objective. Several questions arise concerning the quality of the approximations that result. The behavior of the specific approximations considered in this investigation and the conclusions that derive therefrom are discussed in Section 7.4 and shall not be restated here. In the remaining paragraphs, some general questions relating to the application of the theory and the method of approach are discussed. Topics that require additional investigation and areas for future research are included throughout the discussion.

I. The mathematical model upon which the study is based assumes a time-discrete formulation. Some of the reasons for the use of this model have been discussed in Chapter 1. On the other hand, one major difficulty is created by the use of a time-discrete model that must be overcome before the theory can be applied. In particular, many dynamical systems are described by differential equations rather than difference equations. Although these differential

equations can be solved numerically, it is often impossible to obtain an analytic form for the solution. Thus, the application of the perturbation theory suggested here is complicated by the difficulties inherent in determining the partial derivatives required for the Taylor series expansion of the plant equation. This difficulty is circumvented in using a linear theory by performing the linearization in terms of the differential equations. It is, then, a straightforward matter to establish the state transition matrix and thereby to establish a linear difference equation. It was seen in Chapter 5 that the nonlinear plant effects can be dominant if specialized techniques such as rectification cannot be used. Thus, this aspect cannot be ignored and would appear to require considerable additional research.

II. The relations that are obtained by the application of this nonlinear perturbation theory are considerably more involved than the well-known results for linear systems. Their implementation for use with a multidimensional system would appear to lead to a significant computational burden. Thus, one should examine the possibility of developing special techniques that would enable the utilization of linear theory for problems which at first glance would seem to require more sophisticated methods. This has been shown by example to be possible for estimation problems. It is not clear if analogous policies can be developed for systems involving control considerations that will prove to be as fruitful. In many control problems, the nominal control policy is selected so that prescribed trajectory constraints are satisfied. Thus, one cannot arbitrarily modify the nominal without first verifying that the constraints

will be satisfied by the new nominal. This restriction would seriously hamper the use of a rectification policy. Thus, the development of specialized methods for extending the range of applicability of linear theory for control problems appears to be worthy of consideration.

III. The validity of the approximations has been tested by observing the behavior of the conditional mean (i.e., the minimum mean-square estimate) for specific nonlinear systems. It has been found that the estimates behave satisfactorily in many cases and do provide an improvement relative to the output of a linear filter. However, the approximations were observed to deteriorate quite radically for many problems in which the nonlinear plant and measurement effects were large. The deterioration was marked by the appearance of negative values for the second and fourth central moments. Since this is theoretically impossible, such behavior must be attributed to the inaccuracy of the approximation. The truncated Edgeworth series that were used for the approximation are not true probability density functions because they can assume negative values for some values of the argument. These approximations do not provide the only possibility that could be investigated. It would be interesting to assume that  $p(x_k/z^k)$  belongs to a particular class of parametric distributions (e.g., the Pearson distributions). The parameters would then be determined from the system characteristics. The use of a true density for the approximation might cause the moments to behave more satisfactorily.

IV. Examination of the behavior of the moments provides an indirect method of judging the quality of the approximation. It would be desirable to compare the approximation with the actual a posteriori density. This could be done for a simple system with a static plant and nonlinear measurements corrupted by gaussian noise. That is, consider the system

$$x_k = x \quad \text{for all } k$$

$$z_k = h_k(x) + v_k$$

Let  $x$  and  $v_k$  be gaussian and independent. The a posteriori density can be written as

$$p(x/z^k) = \frac{p(z^k/x)p(x)}{p(z^k)}$$

It is possible to form  $p(z^k/x)$  and  $p(x)$  and

$$p(z^k) = \int p(z^k/x)p(x)dx$$

so the density can be written explicitly.

$$p(z^k/x) = k_k \prod_{i=1}^k \exp - \frac{1}{2} \left[ \frac{z_i - h(x)}{r_i} \right]^2$$

and

$$p(x) = k_x \exp - \frac{1}{2} \left( \frac{x - a}{m_0} \right)^2$$

Thus, one could compute  $p(z^k/x)$  for this simple system and compare it with the results given by the approximation. One could also examine the  $p(z^k/x)$  to determine some asymptotic properties of the density and verify that the approximation exhibits these characteristics.

V. Little consideration has been given to the problem of establishing control policies for nonlinear systems. In Chapter 4, it was suggested that the Separation Principle is not valid for a linear plant when the measurements are nonlinear. The density approximation of Chapter 6 could be used to develop perturbative control laws, and it would be interesting to investigate the policies that result. In fact, these approximations might have their greatest use in the development of nonlinear stochastic control policies. This should provide a fertile area for future investigations.

## REFERENCES

1. Bellman, R. and R. Kalaba, "A Mathematical Theory of Adaptive Control Processes," Proc. Nat. Acad. Sci., Vol. 45, 1959, pp. 1288-1290.
2. Bellman, R., Adaptive Control Processes: A Guided Tour, Princeton University Press, Princeton, New Jersey, 1961.
3. Zadeh, L. A. and C. A. Desoer, Linear System Theory: The State Space Approach, McGraw-Hill, New York, 1963.
4. Fel'dbaum, A. A., "Dual Control Theory - I," Automation and Remote Control, Vol. 21, No. 9, September 1960, pp. 874-880, (translation).
5. Fel'dbaum, A. A., "Dual Control Theory - II," Automation and Remote Control, Vol. 21, No. 11, November 1960, pp. 1033-1039, (translation).
6. Fel'dbaum, A. A., "Dual Control Theory - III," Automation and Remote Control, Vol. 22, No. 1, January 1961, pp. 1-12, (translation).
7. Fel'dbaum, A. A., "Dual Control Theory - IV," Automation and Remote Control, Vol. 22, No. 2, February 1961, pp. 109-121, (translation).
8. Fel'dbaum, A. A., "On the Optimal Control of Markov Objects," Automation and Remote Control, Vol. 23, No. 8, August 1962, pp. 927-941, (translation).
9. Fel'dbaum, A. A., Optimal Control Systems, (Translated by A. Kraiman), Academic Press, New York, 1965.
10. Cramér, H., Mathematical Methods of Statistics, Princeton University Press, Princeton, New Jersey, 1961.
11. Sherman, S., "Non-Mean-Square Error Criteria," IRE Trans. on Information Theory, IT-4, 125-126, 1958.
12. Cox, H., "On the Estimation of State Variables and Parameters for Noisy Dynamic Systems," IEEE Trans. on Auto. Control, AC-9, No. 1, January 1964, pp. 5-12.

13. Sworder, D. D., Synthesis of Optimal, Discrete Time, Adaptive Control Systems, Ph.D. Dissertation, Dept. of Engineering, UCLA, 1964.
14. Sworder, D. D., Optimal Adaptive Control Systems, Academic Press, New York, 1966.
15. Dreyfus, S. E., "Some Types of Optimal Control of Stochastic Systems," SIAM Journal on Control, Series A, Vol. 2, No. 1, 1964.
16. Katz, S., "Best End-point Control of Noisy Systems," J. Electronics and Control, Vol. 12, 1962, pp. 323-343.
17. Simon, H. A., "Dynamic Programming Under Uncertainty with a Quadratic Criterion Function," Econometrica, Vol. 24, 1956, pp. 74-81.
18. Wonham, W. M., "Stochastic Problems in Optimal Control," RIAS Technical Report 63-14, May 1963.
19. Sorenson, H. W., "Some Aspects of Optimal Stochastic Control Theory," AC Electronics Division of General Motors Corporation Technical Report ACRD-LAS-0037, February 1965.
20. Bucy, R. S., "Nonlinear Filtering Theory," IEEE Trans. on Auto. Cont., Vol. AC-10, No. 2, April 1965, p. 198.
21. Bass, R. W., L. Schwartz, and V. D. Norum, "Optimal Multichannel Nonlinear Filtering," Hughes Aircraft Report No. SSD50064R, August 1965.
22. Mortensen, R. E., Optimal Control of Continuous-Time Stochastic Systems, Ph.D. Dissertation, Dept. of Engineering, University of California at Berkeley, 1966.
23. Fisher, J. R., Conditional Probability Density Functions and Optimal Nonlinear Estimation, Ph.D. Dissertation, Dept. of Engineering, UCLA, 1966.
24. Kalman, R. E., "Fundamental Study of Adaptive Control Systems," Tech. Report ASD-TR-61, Vol. 1, 1962.
25. Kalman, R. E., "A New Approach to Linear Filtering and Prediction Problems," J. Basic Engr., Vol. 82D, March 1960, pp. 35-45.

26. Kalman, R. E. and R. Bucy, "New Results in Linear Filtering and Prediction Theory," J. Basic Engr., Vol. 83D, March 1961, pp. 95-108.
27. Kalman, R. E. and R. W. Koepcke, "Optimal Synthesis of Linear Sampling Control Systems Using Generalized Performance Indexes," Trans. ASME, Vol. 80, November 1958, pp. 1820-1826.
28. Gunckel, T. L. and G. F. Franklin, "A General Solution for Linear, Sampled-Data Control," J. Basic Engr., Vol. 85, 1963, pp. 197-203.
29. Joseph, P. and J. T. Tou, "On Linear Control Theory," AIEE Trans. (Appl. and Ind.), Vol. 80, 1961, pp. 193-196.
30. Sorenson, H. W., "Kalman Filtering Techniques," Chapter 6 in Advances in Control System, Volume 3, edited by C. T. Leondes, Academic Press, New York, 1966.
31. Denham, W. F. and S. Pines, "Sequential Estimation When Measurement Function Nonlinearity Is Comparable to Measurement Error," AIAA Journal, Vol. 4, No. 6, June 1966, pp. 1071-1076.
32. Ho, Y. C. and R. C. K. Lee, "A Bayesian Approach to Problems in Stochastic Estimation and Control," Joint Automatic Control Conference, 1964, pp. 382-387.
33. Aoki, M., Optimization of Stochastic Systems: Topics in Discrete-Time Stochastic and Adaptive Systems Optimization, Academic Press, New York, 1966.
34. Stratonovich, R. L., "Conditional Markov Processes," Theory of Probability and Its Application, Vol. 5, 1960, (translation).
35. Lee, R. C. K., Optimal Estimation, Identification, and Control, Research Monograph No. 28, The M.I.T. Press, Cambridge, Mass., 1964.
36. Rauch, H. E., F. Tung, C. T. Striebel, "Maximum Likelihood Estimates of Linear Dynamic Systems," AIAA Journal, Vol. 3, No. 8, August 1965, pp. 1445-1450.
37. Papoulis, A., Probability, Random Variables, and Stochastic Processes, McGraw-Hill, New York, 1965.
38. Doob, J. L., Stochastic Processes, Wiley, 1953.



39. Pugachev, V. S., Theory of Random Functions and Its Application to Control Problems, Addison-Wesley Publishing Company Inc., 1965.
40. Sorenson, H. W., "On the Controllability and Observability of Linear, Stochastic, Time-Discrete Control Systems," AC Electronics Division of General Motors Corporation Technical Report LA-3464, March 1966.
41. Kalman, R. E., Ho, Y. C., Narendra, K. S., "Controllability of Linear Dynamical Systems," Contributions to Differential Equations, Vol. 1, No. 2, pp. 189-213.
42. Deley, G. W., and G. F. Franklin, "Optimal Bounded Control of Linear, Sampled-Data Systems with Quadratic Loss," ASME Journal of Basic Engineering, Vol. 87, Series D, No. 1, March 1965, pp. 135-141.
43. Orford, R. J., Optimization of Stochastic Final-Value Control Systems Subject to Constraints, Ph.D. Dissertation, Polytechnic Institute of Brooklyn, 1962.
44. Gilbert, E. G., "Controllability and Observability in Multivariable Control Systems," SIAM Journal on Control, Series A, Vol. 1, No. 2, 1963, pp. 128-151.
45. Jazwinski, A. H., "Nonlinear Filtering with Discrete Observations," AIAA 3rd Aerospace Sciences Meeting, New York, January 24-26, 1966, AIAA Paper 66-38.
46. Ho, Y. C., "The Method of Least Squares and Optimal Filtering Theory," Rand Corp. Memo. RM-3329-PR, 1962.
47. Wengert, R. E., "A Simple Automatic Derivative Evaluation Program," Communications of the ACM, Vol. 7, No. 8, August 1964, pp. 463-464.
48. Wilkins, R. D., "Investigation of a New Analytical Method for Numerical Derivative Evaluation," Communications of the ACM, Vol. 7, No. 8, August 1964, pp. 465-471.
49. Meditch, J. S., J. L. Lemay, and J. P. Janus, "Analysis of a Horizon Scanner Autonomous Orbital Navigation System," Aerospace Corp. Rept. TDR-469(5540-10)-6, 1965.
50. Sorenson, H. W., Z. Salomon, T. Kido, "Application of Kalman Filter Theory to the Determination of Nearly Circular Orbits" AC Electronics Technical Report, August 1963.

51. Smith, G. L., S. F. Schmidt, L. A. McGee, "Application of Statistical Filter Theory to the Optimal Estimation of Position and Velocity on Board a Circumlunar Vehicle," NASA TR-R-135, 1962.
52. Friedlander, A. L., A Midcourse Guidance Procedure for Electrically Propelled Interplanetary Spacecraft, M.S. Thesis, Case Institute of Technology, 1963.
53. Dusek, H. M., "Theory of Error Compensation in Astro-Inertial Guidance Systems for Low-thrust Space Missions," Progress in Astronautics and Aeronautics, Vol. 13, Academic Press, 1964.
54. Mendelsohn, J., W. O'Dwyer, and A. Kaerchner, "A Numerical Comparison of Two Orbit Determination Methods," Transactions of the Ninth Symposium of Ballistic Missiles, Vol. 1, DDC, 1964.
55. Danby, J. M. A., "Matrix Methods in the Calculation and Analysis of Orbits," AIAA Journal, Vol. 2, No. 1, January 1964.
56. Spragins, J. D. (Jr.), "Reproducing Distributions for Machine Learning," Stanford Electronics Laboratories, Systems Theory Laboratory Technical Report No. 6103-7, November 1963.
57. Stratonovich, R. L., "Topics in the Theory of Random Noise, Volume I," (Revised English Edition, Translated from the Russian by R. A. Silverman), Gordon and Beach Science Publishers, Inc., New York, 1963.
58. Kuznetsov, P. I., R. L. Stratonovich, and V. I. Tikhonov, "Quasi-Moment Functions in the Theory of Random Processes," Theory of Probability and Its Application, Vol. 5, No. 1, 1960, pp. 80-97.
59. Friedman, B., Principles and Techniques of Applied Mathematics, Wiley, 1956.
60. Moulton, F. R., Celestial Mechanics, The Macmillan Company, 15th Printing, 1962.
61. Kahn, H., "Applications of Monte Carlo," Rand Corporation Research Memorandum, RM-1237-AEC, April 1954.
62. Kuznetsov, P. I., R. L. Stratonovich, and V. I. Tikhonov, Nonlinear Transformations of Stochastic Processes (Translation Edited by J. Wise and D. C. Cooper), Pergamon Press, London, 1965.

63. Clarke, V. C. (Jr.), "Constants and Related Data for Use in Trajectory Calculations as Adopted by the Ad Hoc NASA Standard Constant Committee," J. P. L. Technical Report, No. 32-604, March 1964.

## APPENDIX A

### GAUSSIAN A POSTERIORI DENSITY FUNCTION

Let the state be  $n$ -dimensional and suppose that  $m$  measurements are available at each sampling time. The conditional density evolves according to the relationship

$$p(\underline{x}_k / \underline{z}^k) = \frac{p(\underline{x}_k / \underline{z}^{k-1}) p(\underline{z}_k / \underline{x}_k)}{p(\underline{z}_k / \underline{z}^{k-1})} \quad (\text{A. 1})$$

where

$$p(\underline{x}_k / \underline{z}^{k-1}) = \int p(\underline{x}_{k-1} / \underline{z}^{k-1}) p(\underline{x}_k / \underline{x}_{k-1}) d\underline{x}_{k-1}$$

and

$$p(\underline{z}_k / \underline{z}^{k-1}) = \int p(\underline{x}_k / \underline{z}^{k-1}) p(\underline{z}_k / \underline{x}_k) d\underline{x}_k$$

The initial density  $p(\underline{x}_0 / \underline{z}_0)$  is determined from

$$p(\underline{x}_0 / \underline{z}_0) = \frac{p(\underline{z}_0 / \underline{x}_0) p(\underline{x}_0)}{\int p(\underline{z}_0 / \underline{x}_0) p(\underline{x}_0) d\underline{x}_0} \quad (\text{A. 2})$$

The conditions that the plant and measurement equations must satisfy so that  $p(\underline{x}_k / \underline{z}^k)$  can be represented by

$$p(\underline{x}_k / \underline{z}^k) = k_k \exp - \frac{1}{2} \{ (\underline{x}_k - \hat{\underline{x}}_k)^T P_{k-1}^{-1} (\underline{x}_k - \hat{\underline{x}}_k) \} \quad (\text{A. 3})$$

shall be determined.

The initial state and the measurement and plant noise are assumed to have gaussian distributions.

$$p(\underline{x}_0) = k_{x0} \exp - \frac{1}{2} \{ (\underline{x}_0 - \underline{a})^T M_0^{-1} (\underline{x}_0 - \underline{a}) \} \quad (\text{A. 4})$$

$$p(\underline{w}_k) = k_w \exp - \frac{1}{2} \{ \underline{w}_k^T Q_k^{-1} \underline{w}_k \} \quad (A.5)$$

$$p(\underline{v}_k) = k_v \exp - \frac{1}{2} \{ \underline{v}_k^T R_k^{-1} \underline{v}_k \} \quad (A.6)$$

First, determine the density  $p(\underline{x}_0/\underline{z}_0)$ . Since the measurement noise is additive, it follows that

$$p(\underline{z}_0/\underline{x}_0) = k_v \exp - \frac{1}{2} \{ (\underline{z}_0 - \underline{h}_0)^T R_0^{-1} (\underline{z}_0 - \underline{h}_0) \} \quad (A.7)$$

Expand  $\underline{h}_0$  in a Taylor series. It can be written as

$$\underline{h}_0(\underline{x}_0) = \underline{h}_0(\underline{x}_0^*) + H_0 \delta \underline{x}_0 + \frac{1}{2} \eta_0 \quad (A.7)$$

where

$$H_0 \stackrel{\text{Df}}{=} \begin{bmatrix} \frac{\partial h_0^1}{\partial x_0^1} & \dots & \frac{\partial h_0^1}{\partial x_0^n} \\ \vdots & & \vdots \\ \frac{\partial h_0^n}{\partial x_0^1} & \dots & \frac{\partial h_0^n}{\partial x_0^n} \end{bmatrix} = \left( \frac{\partial \underline{h}_0}{\partial \underline{x}} \right)$$

$$\eta_0 \stackrel{\text{Df}}{=} \begin{bmatrix} \delta \underline{x}_0^T J_0^1 \delta \underline{x}_0 \\ \delta \underline{x}_0^T J_0^2 \delta \underline{x}_0 \\ \vdots \\ \delta \underline{x}_0^T J_0^m \delta \underline{x}_0 \end{bmatrix}$$

The  $J_o^i$  are defined as

$$J_o^i \stackrel{\text{Df}}{=} \begin{bmatrix} \frac{\partial^2 h_o^i}{\partial x_1 \partial x_1} & \dots & \frac{\partial^2 h_o^i}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \frac{\partial^2 h_o^i}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 h_o^i}{\partial x_n \partial x_n} \end{bmatrix}$$

The superscript on the  $h_o$  is used to designate the  $i^{\text{th}}$  component of the vector  $\underline{h}_o$ .

Using the expansion in (A.7) and keeping only the quadratic terms, the  $p(\underline{z}_o/\underline{x}_o)$  becomes

$$\begin{aligned} p(\underline{z}_o/\underline{x}_o) = k_v \exp -\frac{1}{2} \{ \delta \underline{z}_o^T R_o^{-1} \delta \underline{z}_o - 2 \delta \underline{z}_o^T R_o^{-1} H_o \delta \underline{x}_o \\ + \delta \underline{x}_o^T H_o^T R_o^{-1} H_o \delta \underline{x}_o - \delta \underline{z}_o^T R_o^{-1} \underline{\eta}_o \} \end{aligned}$$

The term involving  $\underline{\eta}_o$  must be rewritten in terms of  $\delta \underline{x}_o$ . Let

$$\underline{y}_o \stackrel{\text{Df}}{=} R_o^{-1} \delta \underline{z}_o$$

Then

$$\begin{aligned} \underline{\eta}_o^T \underline{y}_o &= [\delta \underline{x}_o^T J_o^1 \delta \underline{x}_o \dots \delta \underline{x}_o^T J_o^m \delta \underline{x}_o] \begin{bmatrix} y_o^1 \\ \vdots \\ y_o^m \end{bmatrix} \\ &= \sum_{i=1}^m \delta \underline{x}_o^T J_o^i \delta \underline{x}_o y_o^i \\ &= \delta \underline{x}_o^T \left( \sum_{i=1}^m J_o^i y_o^i \right) \delta \underline{x}_o \end{aligned}$$

Then,

$$p(\underline{z}_0/\underline{x}_0) = k_v \exp - \frac{1}{2} \{ \delta \underline{z}_0^T R_o^{-1} \delta \underline{z}_0 - 2 \delta \underline{z}_0^T R_o^{-1} H_o \delta \underline{x}_0 + \delta \underline{x}_0^T [H_o^T R_o^{-1} H_o - \sum_{i=1}^m J_o^i y_o^i] \delta \underline{x}_0 \} \quad (A.8)$$

Using (A.4) and (A.8), one obtains

$$p(\underline{z}_0/\underline{x}_0)p(\underline{x}_0) = k_{xo} k_v \exp \{ \delta \underline{z}_0^T R_o^{-1} \delta \underline{z}_0 + \delta \underline{a}^T M_o^{-1} \delta \underline{a} \} \exp - \frac{1}{2} \{ \delta \underline{x}_0^T [H_o^T R_o^{-1} H_o - \sum_{i=1}^m J_o^i y_o^i + M_o^{-1}] \delta \underline{x}_0 - 2 [\delta \underline{a}^T M_o^{-1} + \delta \underline{z}_0^T R_o^{-1} H_o] \delta \underline{x}_0 \}$$

Define

$$P_o^{-1} \stackrel{\text{Df}}{=} H_o^T R_o^{-1} H_o - \sum_{i=1}^m J_o^i y_o^i + M_o^{-1}$$

and

$$\delta \hat{\underline{x}}_o^T \stackrel{\text{Df}}{=} [\delta \underline{z}_0^T R_o^{-1} H_o + \delta \underline{a}^T M_o^{-1}] P_o$$

Using these definitions and completing the square, it follows that

$$p(\underline{z}_0/\underline{x}_0)p(\underline{x}_0) = k_{xo} k_v \exp - \frac{1}{2} \{ \delta \underline{z}_0^T R_o^{-1} \delta \underline{z}_0 + \delta \underline{a}^T M_o^{-1} \delta \underline{a} - \delta \hat{\underline{x}}_o^T P_o^{-1} \delta \hat{\underline{x}}_o \} \exp - \frac{1}{2} \{ (\delta \underline{x}_0 - \delta \hat{\underline{x}}_o)^T P_o^{-1} (\delta \underline{x}_0 - \delta \hat{\underline{x}}_o) \} \quad (A.9)$$

Integrate with respect to  $\delta \underline{x}_0$ . Then,

$$\int p(\underline{z}_0/\underline{x}_0)p(\underline{x}_0) d\underline{x}_0 = \frac{k_{xo} k_v}{k_o} \exp - \frac{1}{2} \{ \delta \underline{z}_0^T R_o^{-1} \delta \underline{z}_0 + \delta \underline{a}^T M_o^{-1} \delta \underline{a} - \delta \hat{\underline{x}}_o^T R_o^{-1} \delta \hat{\underline{x}}_o \} \quad (A.10)$$

where  $k_0$  is the normalization constant.

$$k_0 = [(2\pi)^n (\det P_0)]^{-1/2}$$

From (A.2), (A.9) and (A.10), one sees that

$$p(\underline{x}_0/\underline{z}_0) = k_0 \exp - \frac{1}{2} \{ \delta \underline{x}_0 - \delta \hat{\underline{x}}_0 \}^T P_0^{-1} \{ \delta \underline{x}_0 - \delta \hat{\underline{x}}_0 \} \quad (A.11)$$

This can be rewritten as

$$p(\underline{x}_0/\underline{z}_0) = k_0 \exp - \frac{1}{2} \{ (\underline{x}_0 - \hat{\underline{x}}_0)^T P_0^{-1} (\underline{x}_0 - \hat{\underline{x}}_0) \} \quad (A.12)$$

where

$$\hat{\underline{x}}_0 = \underline{x}_0^* + \delta \hat{\underline{x}}_0$$

The density has the desired form at  $t_0$ .

Assume at  $t_k$  that

$$p(\underline{x}_k/\underline{z}^k) = k_k \exp - \frac{1}{2} \{ \underline{x}_k - \hat{\underline{x}}_k \}^T P_k^{-1} \{ \underline{x}_k - \hat{\underline{x}}_k \} \quad (A.13)$$

Now, derive  $p(\underline{x}_{k+1}/\underline{z}^{k+1})$ . First, form

$$p(\underline{x}_{k+1}/\underline{x}_k) = k_w \exp - \frac{1}{2} \{ (\underline{x}_{k+1} - \underline{f}_{k+1})^T Q_k^{-1} (\underline{x}_{k+1} - \underline{f}_{k+1}) \} \quad (A.14)$$

Assume a nominal  $\underline{x}_k^*$  and let

$$\underline{f}_{k+1} = \underline{f}_{k+1}(\underline{x}_k^*)$$

Expand  $\underline{f}_{k+1}$  in a Taylor series about  $\underline{x}_k^*$  and retain only second order terms

in  $\delta \underline{x}_k$



$$\begin{aligned}
& (\underline{x}_{k+1} - \underline{f}_{k+1})^T Q_k^{-1} (\underline{x}_{k+1} - \underline{f}_{k+1}) = \\
& [\underline{x}_{k+1} - \underline{f}_{k+1}(\underline{x}_k^*) - F_k \delta \underline{x}_k - \frac{1}{2} \varphi_k (\delta \underline{x}_k \delta \underline{x}_k^T)] Q_k^{-1} (\underline{x}_{k+1} - \underline{f}_{k+1}) \\
& = \delta \underline{x}_{k+1}^T Q_k^{-1} \delta \underline{x}_{k+1} - 2 \delta \underline{x}_{k+1}^T Q_k^{-1} F_k \delta \underline{x}_k + \delta \underline{x}_k^T F_k^T Q_k^{-1} F_k \delta \underline{x}_k \\
& \quad - \delta \underline{x}_{k+1}^T Q_k^{-1} \varphi_k
\end{aligned} \tag{A.15}$$

where

$$\begin{aligned}
\delta \underline{x}_{k+1} & \stackrel{\text{Df}}{=} \underline{x}_{k+1} - \underline{f}_{k+1}(\underline{x}_k^*) \\
\delta \underline{x}_k & \stackrel{\text{Df}}{=} \underline{x}_k - \underline{x}_k^* \\
F_k & \stackrel{\text{Df}}{=} \left( \frac{\partial \underline{f}_{k+1}}{\partial \underline{x}_k} \right) = \begin{bmatrix} \frac{\partial f_{k+1}^1}{\partial x_k^1} & \dots & \frac{\partial f_{k+1}^1}{\partial x_k^n} \\ \vdots & & \vdots \\ \frac{\partial f_{k+1}^n}{\partial x_k^1} & \dots & \frac{\partial f_{k+1}^n}{\partial x_k^n} \end{bmatrix}
\end{aligned}$$

and

$$\begin{aligned}
\varphi_k & \stackrel{\text{Df}}{=} \begin{bmatrix} \delta \underline{x}_k^T & G_k^1 & \delta \underline{x}_k \\ \vdots & \vdots & \vdots \\ \delta \underline{x}_k^T & G_k^n & \delta \underline{x}_k \end{bmatrix} \\
G_k^i & \stackrel{\text{Df}}{=} \left( \frac{\partial^2 f_{k+1}^i}{\partial x_k^l \partial x_k^j} \right)
\end{aligned}$$

For the moment, let

$$\underline{\omega} = Q_k^{-1} \delta \underline{x}_{k+1}$$

Then, it can easily be shown that

$$\delta \underline{x}_{k+1}^T Q_k^{-1} \underline{\omega} = \delta \underline{x}_k^T \left( \sum_{i=1}^n G_k^i \omega^i \right) \delta \underline{x}_k$$

Using (A.13) through (A.15), we obtain

$$\begin{aligned} p(\underline{x}_k / \underline{z}^k) p(\underline{x}_{k+1} / \underline{x}_k) &= k_k k_w \exp - \frac{1}{2} \{ \delta \underline{x}_{k+1}^T Q_k^{-1} \delta \underline{x}_{k+1} + \delta \hat{\underline{x}}_k^T P_k^{-1} \delta \hat{\underline{x}}_k \} \\ &\exp - \frac{1}{2} \{ \delta \underline{x}_k^T [F_k^T Q_k^{-1} F_k - \sum_{i=1}^n G_k^i \omega^i + P_k^{-1}] \delta \underline{x}_k \\ &\quad - 2[\delta \underline{x}_{k+1}^T Q_k^{-1} F_k + \delta \hat{\underline{x}}_k^T P_k^{-1}] \delta \underline{x}_k \} \end{aligned} \quad (A.16)$$

Let

$$\begin{aligned} B_k^{-1} &\stackrel{\text{Df}}{=} F_k^T Q_k^{-1} F_k - \sum_{i=1}^n G_k^i \omega^i + P_k^{-1} \\ \delta \underline{v}_k^T &\stackrel{\text{Df}}{=} [\delta \underline{x}_{k+1}^T Q_k^{-1} F_k + \delta \hat{\underline{x}}_k^T P_k^{-1}] B_k \end{aligned}$$

By completing the square, (A.16) becomes

$$\begin{aligned} p(\underline{x}_k / \underline{z}^k) p(\underline{x}_{k+1} / \underline{x}_k) &= k_k k_w \\ &\exp - \frac{1}{2} \{ \delta \underline{x}_{k+1}^T Q_k^{-1} \delta \underline{x}_{k+1} + \delta \hat{\underline{x}}_k^T P_k^{-1} \delta \hat{\underline{x}}_k - \delta \underline{v}_k^T B_k^{-1} \delta \underline{v}_k \} \\ &\exp - \frac{1}{2} \{ (\delta \underline{x}_k - \delta \underline{v}_k)^T B_k^{-1} (\delta \underline{x}_k - \delta \underline{v}_k) \} \end{aligned}$$

Integrate with respect to  $\underline{x}_k$

$$p(\underline{x}_{k+1}/\underline{z}^k) = \frac{k_k^k w}{k_{Nk}} \exp - \frac{1}{2} \{ \delta \underline{x}_{k+1}^T Q_k^{-1} \delta \underline{x}_{k+1} + \delta \hat{\underline{x}}_k^T P_k^{-1} \delta \hat{\underline{x}}_k - \delta \underline{v}_k^T B_k^{-1} \delta \underline{v}_k \} \quad (A.17)$$

where the normalization constant  $k_{Nk}$  is

$$k_{Nk} = [(2\pi)^n | B_k | ]^{-1/2}$$

The  $\delta \underline{v}_k^T B_k^{-1} \delta \underline{v}_k$  contains  $\delta \underline{x}_{k+1}$ , so this term must be approximated by a quadratic.

$$\begin{aligned} B_k &= [F_k^T Q_{k-1}^{-1} F_k + P_k^{-1} - \sum_{i=1}^n G_k^i \omega^i]^{-1} \\ &= [I - (F_k^T Q_{k-1}^{-1} F_k + P_k^{-1})^{-1} \sum_{i=1}^n G_k^i \omega^i]^{-1} (F_k^T Q_{k-1}^{-1} F_k + P_k^{-1})^{-1} \end{aligned}$$

Let

$$\Pi_{k+1} \stackrel{\text{Df}}{=} F_k^T Q_{k-1}^{-1} F_k + P_k^{-1} \quad (A.18)$$

so

$$B_k \stackrel{\text{Df}}{=} [I - \Pi_{k+1}^{-1} \sum_{i=1}^n G_k^i \omega^i]^{-1} \Pi_{k+1}^{-1}$$

Assuming that  $\Pi_{k+1}^{-1} (\sum_{i=1}^n G_k^i \omega^i)$  has sufficiently small norm, the Neumann series expansion [59] is approximately

$$B_k = \Pi_{k+1}^{-1} + \Pi_{k+1}^{-1} \left( \sum_{i=1}^n G_k^i \omega^i \right) \Pi_{k+1}^{-1} + \Pi_{k+1}^{-1} \left( \sum_{i=1}^n G_k^i \omega^i \right) \Pi_{k+1}^{-1} \left( \sum_{i=1}^n G_k^i \omega^i \right) \Pi_{k+1}^{-1}$$

$$\begin{aligned}
\delta \underline{v}_k^T B_k^{-1} \delta \underline{v}_k &= (\delta \underline{x}_{k+1}^T Q_k^{-1} F_k + \delta \hat{\underline{x}}_k^T P_k^{-1}) B_k (F_k^T Q_k^{-1} \delta \underline{x}_{k+1} + P_k^{-1} \delta \hat{\underline{x}}_k) \\
&= (\delta \underline{x}_{k+1}^T Q_k^{-1} F_k + \delta \hat{\underline{x}}_k^T P_k^{-1}) \Pi_{k+1}^{-1} (F_k^T Q_k^{-1} \delta \underline{x}_{k+1} + P_k^{-1} \delta \hat{\underline{x}}_k) \\
&\quad + \delta \hat{\underline{x}}_k^T P_k^{-1} \Pi_{k+1}^{-1} \left( \sum_{i=1}^n G_k^i \omega^i \right) \Pi_{k+1}^{-1} P_k^{-1} \delta \hat{\underline{x}}_k \\
&\quad + 2 \delta \hat{\underline{x}}_k^T P_k^{-1} \Pi_{k+1}^{-1} \left( \sum_{i=1}^n G_k^i \omega^i \right) \Pi_{k+1}^{-1} F_k^T Q_k^{-1} \delta \underline{x}_{k+1} \\
&\quad + \delta \hat{\underline{x}}_k^T P_k^{-1} \Pi_{k+1}^{-1} \left( \sum_{i=1}^n G_k^i \omega^i \right) \Pi_{k+1}^{-1} \left( \sum_{i=1}^n G_k^i \omega^i \right) \Pi_{k+1}^{-1} P_k^{-1} \delta \hat{\underline{x}}_k
\end{aligned}$$

where all terms of greater than second order have been neglected.

Consider  $\left( \sum_{i=1}^n G_k^i \omega^i \right) \Pi_{k+1}^{-1} P_k^{-1} \delta \hat{\underline{x}}_k$ . The  $\omega^i$  are scalar quantities, so

$$\left( \sum_{i=1}^n G_k^i \omega^i \right) \Pi_{k+1}^{-1} P_k^{-1} \delta \hat{\underline{x}}_k = \sum_{i=1}^n G_k^i \Pi_{k+1}^{-1} P_k^{-1} \delta \hat{\underline{x}}_k \omega^i$$

But  $\underline{\omega} = Q_k^{-1} \delta \underline{x}_{k+1}$ . Denote the  $i^{\text{th}}$  row of  $Q_k^{-1}$  by  $q_i^T$ . Then

$$\omega^i = q_i^T \delta \underline{x}_{k+1}$$

$$\sum_{i=1}^n G_k^i \omega^i \Pi_{k+1}^{-1} P_k^{-1} \delta \hat{\underline{x}}_k = \left( \sum_{i=1}^n G_k^i \Pi_{k+1}^{-1} P_k^{-1} \delta \hat{\underline{x}}_k q_i^T \right) \delta \underline{x}_{k+1}$$

Let

$$E_k = \sum_{i=1}^n G_k^i \Pi_{k+1}^{-1} P_k^{-1} \delta \hat{\underline{x}}_k q_i^T$$

Then

$$\begin{aligned}
\delta \underline{v}_k^T B_k^{-1} \delta \underline{v}_k &= \delta \hat{\underline{x}}_k^T P_k^{-1} \Pi_{k+1}^{-1} P_k^{-1} \delta \hat{\underline{x}}_k \\
&+ [2\delta \hat{\underline{x}}_k^T P_k^{-1} \Pi_{k+1}^{-1} F_k^T Q_k^{-1} + \delta \hat{\underline{x}}_k^T P_k^{-1} \Pi_{k+1}^{-1} E_k] \delta \underline{x}_{k+1} \\
&+ \delta \underline{x}_{k+1}^T [Q_k^{-1} F_k \Pi_{k+1}^{-1} F_k^T Q_k^{-1} + 2E_k^T \Pi_{k+1}^{-1} F_k Q_k^{-1} + E_k^T \Pi_{k+1}^{-1} E_k] \delta \underline{x}_{k+1} \quad (A.19)
\end{aligned}$$

The introduction of (A.19) into (A.17) gives

$$\begin{aligned}
p(\underline{x}_{k+1}/\underline{z}^k) &= \frac{k_k^k w}{k_{Nk}} \exp - \frac{1}{2} \{ \delta \hat{\underline{x}}_k^T (P_k^{-1} - P_k^{-1} \Pi_{k+1}^{-1} P_k^{-1}) \delta \hat{\underline{x}}_k \} \\
&\exp - \frac{1}{2} \{ \delta \underline{x}_{k+1}^T [Q_k^{-1} - Q_k^{-1} F_k \Pi_{k+1}^{-1} F_k^T Q_k^{-1} - 2E_k^T \Pi_{k+1}^{-1} F_k Q_k^{-1} - E_k^T \Pi_{k+1}^{-1} E_k] \delta \underline{x}_{k+1} \\
&- 2\delta \hat{\underline{x}}_k^T P_k^{-1} \Pi_{k+1}^{-1} [F_k Q_k^{-1} + \frac{1}{2} E_k] \delta \underline{x}_{k+1} \} \quad (A.20)
\end{aligned}$$

The density  $p(\underline{z}_{k+1}/\underline{x}_{k+1})$  is

$$p(\underline{z}_{k+1}/\underline{x}_{k+1}) = k_v \exp - \frac{1}{2} \{ (\underline{z}_{k+1} - \underline{h}_{k+1})^T R_{k+1}^{-1} (\underline{z}_{k+1} - \underline{h}_{k+1}) \}$$

Following the procedure used for  $p(\underline{z}_0/\underline{x}_0)$ , this becomes

$$\begin{aligned}
p(\underline{z}_{k+1}/\underline{x}_{k+1}) &= k_v \exp - \frac{1}{2} \{ \delta \underline{z}_{k+1}^T R_{k+1}^{-1} \delta \underline{z}_{k+1} \} \\
&\exp - \frac{1}{2} \{ \delta \underline{x}_{k+1}^T [H_{k+1}^T R_{k+1}^{-1} H_{k+1} - \sum_{i=1}^m J_{k+1}^i Y_{k+1}^i] \delta \underline{x}_{k+1} \\
&- 2\delta \underline{z}_{k+1}^T R_{k+1}^{-1} H_{k+1} \delta \underline{x}_{k+1} \} \quad (A.21)
\end{aligned}$$

Then, one sees that

$$\begin{aligned}
p(\underline{x}_{k+1}/\underline{z}^k)p(\underline{z}_{k+1}/\underline{x}_{k+1}) &= \frac{k_k^k w_k^k v_k}{k_{Nk}} \exp - \frac{1}{2} \{ \delta \underline{z}_{k+1}^T R_{k+1}^{-1} \delta \underline{z}_{k+1} \\
&\quad + \delta \hat{\underline{x}}_k^T (P_k^{-1} - P_k^{-1} \Pi_{k+1}^{-1} P_k^{-1}) \delta \hat{\underline{x}}_k \} \\
&\quad \exp - \frac{1}{2} \{ \delta \underline{x}_{k+1}^T P_{k+1}^{-1} \delta \underline{x}_{k+1} - 2 \delta \hat{\underline{x}}_{k+1}^T P_{k+1}^{-1} \delta \underline{x}_{k+1} \} \quad (A.22)
\end{aligned}$$

where

$$\begin{aligned}
P_{k+1}^{-1} &\stackrel{\text{Df}}{=} Q_k^{-1} - Q_k^{-1} F_k \Pi_{k+1}^{-1} F_k^T Q_k^{-1} - 2 E_k^T \Pi_{k+1}^{-1} F_k Q_k^{-1} - E_k^T \Pi_{k+1}^{-1} E_k \\
&\quad + H_{k+1}^T R_{k+1}^{-1} H_{k+1} - \sum_{i=1}^m J_{k+1}^i y_{k+1}^i
\end{aligned}$$

$$\delta \hat{\underline{x}}_{k+1}^T \stackrel{\text{Df}}{=} [\delta \hat{\underline{x}}_k^T P_k^{-1} \Pi_{k+1}^{-1} (F_k Q_k^{-1} + \frac{1}{2} E_k) + \delta \underline{z}_{k+1}^T R_{k+1}^{-1} H_{k+1}] P_{k+1}$$

Completing the square and integrating with respect to  $\delta \underline{x}_{k+1}$  gives

$$\begin{aligned}
p(\underline{z}_{k+1}/\underline{z}^k) &= \frac{k_k^k w_k^k v_k}{k_{k+1} k_{Nk}} \exp - \frac{1}{2} \{ \delta \underline{z}_{k+1}^T R_{k+1}^{-1} \delta \underline{z}_{k+1} \\
&\quad + \delta \hat{\underline{x}}_k^T (P_k^{-1} - P_k^{-1} \Pi_{k+1}^{-1} P_k^{-1}) \delta \hat{\underline{x}}_k - \delta \hat{\underline{x}}_{k+1}^T P_{k+1}^{-1} \delta \hat{\underline{x}}_{k+1} \} \quad (A.23)
\end{aligned}$$

where

$$k_{k+1} = [(2\pi)^n | P_{k+1} | ]^{-1/2}$$

Division of (A.22) by (A.23) yields  $p(\underline{x}_{k+1}/\underline{z}^{k+1})$ .

$$\begin{aligned}
p(\underline{x}_{k+1}/\underline{z}^{k+1}) &= k_{k+1} \exp - \frac{1}{2} \{ (\delta \underline{x}_{k+1} - \delta \hat{\underline{x}}_{k+1})^T P_{k+1}^{-1} (\delta \underline{x}_{k+1} - \delta \hat{\underline{x}}_{k+1}) \} \\
&= k_{k+1} \exp - \frac{1}{2} \{ (\underline{x}_{k+1} - \hat{\underline{x}}_{k+1})^T P_{k+1}^{-1} (\underline{x}_{k+1} - \hat{\underline{x}}_{k+1}) \} \quad (A.24)
\end{aligned}$$

## APPENDIX B

### MATHEMATICAL MODEL FOR THE SPACE NAVIGATION PROBLEM

The digital computer program utilized for the study presented in Chapter 5 has been developed by the AC Electronics Division of General Motors Corporation to investigate the general problem of guidance and navigation for interplanetary space vehicles. In this appendix, the general equations that are relevant to the problem of Chapter 5 shall be stated with a minimum of accompanying discussion.

The motion of the spacecraft has been assumed to occur about a single central body as described by (5.1), so analytical solutions of the equations are possible. Thus, the position and velocity at any sampling time are obtained from explicit analytical expressions rather than by numerical integration. These expressions are well-known [60].

For this model, it is possible to obtain explicit expressions for the state transition matrix appearing in (5.6) [55].

Let

$$\Phi \stackrel{\text{Def}}{=} \begin{bmatrix} \Phi_1 & \Phi_2 \\ \Phi_3 & \Phi_4 \end{bmatrix} \quad (\text{B. 1})$$

When the nominal is contained in the X-Y plane, the  $\Phi_i$  have the form

$$\Phi_1 = \begin{bmatrix} \Phi_{11} & \Phi_{12} & 0 \\ \Phi_{21} & \Phi_{22} & 0 \\ 0 & 0 & \Phi_{33} \end{bmatrix} ; \quad \Phi_2 = \begin{bmatrix} \Phi_{14} & \Phi_{15} & 0 \\ \Phi_{24} & \Phi_{25} & 0 \\ 0 & 0 & \Phi_{36} \end{bmatrix}$$

$$\Phi_3 = \begin{bmatrix} \Phi_{41} & \Phi_{42} & 0 \\ \Phi_{51} & \Phi_{52} & 0 \\ 0 & 0 & \Phi_{63} \end{bmatrix} ; \quad \Phi_4 = \begin{bmatrix} \Phi_{44} & \Phi_{45} & 0 \\ \Phi_{54} & \Phi_{55} & 0 \\ 0 & 0 & \Phi_{66} \end{bmatrix}$$

Let the eccentric anomaly at  $t_k$  be  $E_k$  and let the eccentricity and mean angular rate be  $e$  and  $n$ . Define  $S_k$  and  $C_k$  as

$$\sin E_k = S_k ; \quad \cos E_k = C_k$$

Then, the elements of the submatrices are

$$\Phi_{11} = \frac{1}{(1-e)^2(1-e C_k)} [C_k^2(1+e-e^2) + C_k(2+e+2e^2-e^3) - 2 - 5e + 2e^2 + 3E_k S_k]$$

$$\Phi_{12} = \frac{\sqrt{1-e^2}}{(1-e)(1-e C_k)} S_k (1-C_k)$$

$$\Phi_{21} = \frac{\sqrt{1-e^2}}{(1-e)^2(1-e C_k)} [S_k C_k(1+e) + S_k(2-e) - 3E_k C_k]$$

$$\Phi_{22} = \frac{1}{(1-e)(1-e C_k)} [C_k^2 + C_k(-1-2e+e^2) + 1]$$

$$\Phi_{33} = \frac{(C_k - e)}{(1-e)}$$

$$\Phi_{14} = \frac{(1-e)}{n(1-e C_k)} S_k [-C_k(1+e) + 2]$$

$$\Phi_{15} = \frac{\sqrt{1-e^2}}{(1-e)n(1-e C_k)} [C_k^2(2-e) + 2C_k(1+e) - 4 - e + 3E_k S_k]$$



$$\Phi_{24} = \frac{\sqrt{1-e^2}}{n(1-e C_k)} [1 - C_k]^2$$

$$\Phi_{25} = \frac{1}{n(1-e C_k)} [S_k C_k (2+e+e^2) + 2S_k - 3(1+e)E_k C_k]$$

$$\Phi_{36} = \frac{S_k (1-e)}{n}$$

$$\Phi_{41} = \frac{n}{(1-e)^2 (1-e C_k)^3} [S_k C_k^2 (e+e^2-e^3) + S_k C_k (-2-5e+2e^2) + S_k (1+e+3e^2-e^3) + 3E_k (C_k-e)]$$

$$\Phi_{42} = \frac{n\sqrt{1-e^2}}{(1-e)(1-e C_k)^3} [e C_k^3 - 2C_k^2 + C_k + 1 - e]$$

$$\Phi_{51} = \frac{n\sqrt{1-e^2}}{(1-e)^2 (1-e C_k)^3} [-C_k^3 (e+e^2) + C_k^2 (2+5e) - C_k (1+e) - 1 - 3e + e^2 + 3E_k S_k]$$

$$\Phi_{52} = \frac{n S_k}{(1-e)(1-e C_k)^3} [e C_k^2 - 2C_k + 1 + e - e^2]$$

$$\Phi_{63} = \frac{-n S_k}{(1-e)(1-e C_k)}$$

$$\Phi_{44} = \frac{1-e}{(1-e C_k)^3} [C_k^3 (e+e^2) - 2C_k^2 (1+e) + 2C_k + 1 - e]$$

$$\Phi_{45} = \frac{\sqrt{1-e^2}}{(1-e)(1-e C_k)^3} [S_k C_k^2 (2e-e^2) - S_k C_k (4+e) + S_k (1+e)^2 + 3E_k (C_k-e)]$$

$$\Phi_{54} = \frac{S_k \sqrt{1-e^2}}{(1-e C_k)^3} [e C_k^2 - 2C_k + 2 - e]$$

$$\Phi_{55} = \frac{1}{(1-e C_k)^3} [-C_k^3(2e+e^2+e^3) + C_k^2(4+5e+5e^2) - C_k(1+3e) - 2 - 3e - e^2 + 3(1+e)E_k S_k]$$

$$\Phi_{66} = \frac{(1-e) C_k}{(1-e C_k)}$$

The computation of the transition matrix is carried out in an in-plane cartesian coordinate system and then transformed using a rotation matrix into the basic nonrotating cartesian system. The evaluation of the two body equations and the transition matrix equations provides the nominal and actual trajectories and the linear model.

A horizon sensor has been assumed to be available to provide data for navigation purposes. This instrument provides a measurement of the direction of the line of sight to the center of a reference body relative to a nonrotating reference frame and a measurement of the angle subtended by the reference body. These angles are depicted in Figure B-1.

These measurements are described by the three angles,  $\alpha$ ,  $\delta$ , and  $\beta$ . The local vertical is defined by the elevation angle  $\alpha$  and the azimuth angle  $\delta$ . The  $\alpha$  is defined to be positive when the vehicle is below the X-Y plane and  $\delta$  is measured counterclockwise from the X axis.

$$\alpha = -\sin^{-1} \left( \frac{X_3}{R} \right) \quad (B-2)$$

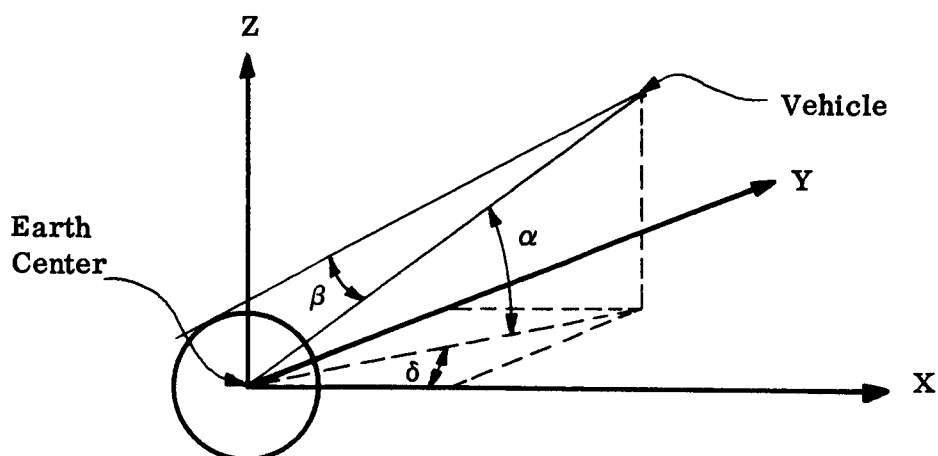


Figure B-1. Geometry of Angular Measurements

$$\delta = \begin{cases} \sin^{-1} \frac{X_2}{(X_1^2 + X_2^2)^{1/2}} \\ \cos^{-1} \frac{X_1}{(X_1^2 + X_2^2)^{1/2}} \end{cases} \quad (\text{B-3})$$

where

$$R = (X_1^2 + X_2^2 + X_3^2)^{1/2}$$

The subtended angle  $\beta$  is given by

$$\beta = \sin^{-1} \frac{r_o}{R} \quad (\text{B-4})$$

where

$$r_o = \text{radius of reference body.}$$

Both the first and second order partial derivatives are required for the navigation procedure. The first order observation matrix  $H_k$  is, in general,

$$H_k \triangleq \left( \frac{\partial h_k}{\partial \underline{X}_k} \right)$$

More specifically, this can be written as

$$H_k = \begin{bmatrix} \frac{\partial \alpha}{\partial X_1} & \frac{\partial \alpha}{\partial X_2} & \frac{\partial \alpha}{\partial X_3} & 0 & 0 & 0 \\ \frac{\partial \delta}{\partial X_1} & \frac{\partial \delta}{\partial X_2} & \frac{\partial \delta}{\partial X_3} & 0 & 0 & 0 \\ \frac{\partial \beta}{\partial X_1} & \frac{\partial \beta}{\partial X_2} & \frac{\partial \beta}{\partial X_3} & 0 & 0 & 0 \end{bmatrix} \quad (B-5)$$

where the derivatives taken with respect to the velocity components of the state vector (i.e.,  $X_4, X_5, X_6$ ) are identically zero as indicated. Denote the non-zero submatrix as  $H_1(t_k)$ . It follows that

$$H_1(t_k) = \begin{bmatrix} \frac{-\sin \alpha \cos \delta}{R} & \frac{-\sin \alpha \sin \delta}{R} & \frac{-(X_1^2 + X_2^2)^{1/2}}{R^2} \\ \frac{-\sin^2 \delta}{X_2} & \frac{\cos^2 \delta}{X_1} & 0 \\ \frac{-X_1 \tan \beta}{R^2} & \frac{-X_2 \tan \beta}{R^2} & \frac{-X_3 \tan \beta}{R^2} \end{bmatrix} \quad (B-6)$$

Next, let

$$J_k^1 \triangleq \left( \frac{\partial^2 \alpha}{\partial \underline{X}^2} \right) ; \quad J_k^2 \triangleq \left( \frac{\partial^2 \delta}{\partial \underline{X}^2} \right) ; \quad J_k^3 \triangleq \left( \frac{\partial^2 \beta}{\partial \underline{X}^2} \right)$$

where

$$\left( \frac{\partial^2 y}{\partial \underline{X}^2} \right) \stackrel{\text{Df}}{=} \begin{bmatrix} \frac{\partial^2 y}{\partial X_1^2} & \frac{\partial^2 y}{\partial X_2 \partial X_1} & \dots & \frac{\partial^2 y}{\partial X_6 \partial X_1} \\ \frac{\partial^2 y}{\partial X_1 \partial X_2} & \frac{\partial^2 y}{\partial X_2^2} & \dots & \frac{\partial^2 y}{\partial X_6 \partial X_2} \\ \frac{\partial^2 y}{\partial X_1 \partial X_6} & \dots & \dots & \frac{\partial^2 y}{\partial X_6^2} \end{bmatrix}$$

There is no dependence upon velocity, so the partial derivatives relative to velocity components are identically zero. Denote the upper left-hand (3x3) submatrix by  $J_1^i(t_k)$  ( $i = 1, 2, 3$ ). Then,

$$J_k^1(t_k) = \begin{bmatrix} J_{11}^1 & J_{12}^1 & J_{13}^1 \\ J_{21}^1 & J_{22}^1 & J_{23}^1 \\ J_{31}^1 & J_{32}^1 & J_{33}^1 \end{bmatrix} \quad (\text{B-7})$$

$$J_{11}^1 = \frac{-\sin \alpha}{R(X_1^2 + X_2^2)^{1/2}} \left[ \sin^2 \delta - 2 \frac{X_1^2}{R^2} \right]$$

$$J_{12}^1 = \frac{X_2 \sin \alpha \cos \delta}{R} \left[ \frac{2}{R^2} + \frac{\sin^2 \delta}{X_2^2} \right]$$

$$J_{13}^1 = \frac{\cos \delta}{R^2} (1 - 2 \sin^2 \alpha)$$

$$J_{21}^1 = \frac{X_2 \sin \alpha \cos \delta}{R} \left[ \frac{2}{R^2} + \frac{\sin^2 \delta}{X_2^2} \right]$$

$$J_{22}^1 = \frac{-\sin \alpha}{R(X_1^2 + X_2^2)^{1/2}} (\cos^2 \delta - \frac{2X_2^2}{R^2})$$

$$J_{23}^1 = \frac{\sin \delta}{R^2} (1 - 2 \sin^2 \alpha)$$

$$J_{31}^1 = \frac{\cos \delta}{R^2} (1 - 2 \sin^2 \alpha)$$

$$J_{32}^1 = \frac{\sin \delta}{R^2} (1 - 2 \sin^2 \alpha)$$

$$J_{33}^1 = \frac{-2 \sin \alpha (X_1^2 + X_2^2)^{1/2}}{R^3}$$

$$J_k^2(t_k) = \begin{bmatrix} J_{11}^2 & J_{12}^2 & J_{13}^2 \\ J_{21}^2 & J_{22}^2 & J_{23}^2 \\ J_{31}^2 & J_{32}^2 & J_{33}^2 \end{bmatrix}$$

(B-8)

$$J_{11}^2 = \frac{2X_1 X_2}{(X_1^2 + X_2^2)^2}$$

$$J_{12}^2 = \frac{-1}{X_1^2 + X_2^2} (1 - 2 \sin^2 \delta)$$

$$J_{21}^2 = \frac{-1}{X_1^2 + X_2^2} (1 - 2 \sin^2 \delta)$$

$$J_{22}^2 = \frac{-2X_1 X_2}{(X_1^2 + X_2^2)^2}$$

$$J_{31}^2 = 0 = J_{32}^2 = J_{33}^2 = J_{13}^2 = J_{23}^2$$

$$J_{k(t_k)}^3 = \begin{bmatrix} J_{11}^3 & J_{12}^3 & J_{13}^3 \\ J_{21}^3 & J_{22}^3 & J_{23}^3 \\ J_{31}^3 & J_{32}^3 & J_{33}^3 \end{bmatrix}$$

(B-9)

$$J_{11}^3 = \frac{-\tan \beta}{R^2} \left( 1 - \frac{X_1^2 \tan^2 \beta}{r_o^2} - \frac{2X_1^2}{R^2} \right)$$

$$J_{12}^3 = \frac{X_1 X_2 \tan \beta}{R^2} \left( \frac{\tan^2 \beta}{r_o^2} + \frac{2}{R^2} \right)$$

$$J_{13}^3 = \frac{X_1 X_3 \tan \beta}{R^2} \left( \frac{\tan^2 \beta}{r_o^2} + \frac{2}{R^2} \right)$$

$$J_{21}^3 = \frac{X_1 X_2 \tan \beta}{R^2} \left( \frac{\tan^2 \beta}{r_o^2} + \frac{2}{R^2} \right)$$

$$J_{22}^3 = \frac{-\tan \beta}{R^2} \left( 1 - \frac{X_2^2 \tan^2 \beta}{r_o^2} - \frac{2X_2^2}{R^2} \right)$$

$$J_{23}^3 = \frac{X_2 X_3 \tan \beta}{R^2} \left( \frac{2}{R^2} + \frac{\tan^2 \beta}{r_o^2} \right)$$

$$J_{31}^3 = \frac{X_1 X_3 \tan \beta}{R^2} \left( \frac{\tan^2 \beta}{r_o^2} + \frac{2}{R^2} \right)$$

$$J_{32}^3 = \frac{X_2 X_3 \tan \beta}{R^2} \left( \frac{2}{R^2} + \frac{\tan^2 \beta}{r_o^2} \right)$$

$$J_{33}^3 = \frac{-\tan \beta}{R^2} \left( 1 - \frac{X_3^2 \tan^2 \beta}{r_o^2} - \frac{2X_3^2}{R^2} \right)$$

Note that the  $J_1^i(t_k)$  are symmetric.

The preceding equations provide the complete model for the system. To accomplish the Monte Carlo simulation, the actual trajectory of the spacecraft is computed in addition to the nominal trajectory. The initial deviation of the ensemble with mean zero and covariance matrix  $M_o$  (see (3.3) ) is selected using a gaussian random number generator [61].

It is the state of the actual trajectory that is to be estimated. The measurement data are computed from the nominal and actual trajectories. The exact measurement values are corrupted by adding at each time numbers from a gaussian random number generator assuming the ensemble has mean zero and covariance matrix  $R_k$ .



## APPENDIX C

### GRAM-CHARLIER AND EDGEWORTH EXPANSIONS

The Gram-Charlier and Edgeworth [10, 23, 39] expansions have been utilized in Chapter 6 to approximate the a posteriori density function  $p(x_k/z^k)$ . The purpose of this appendix is to discuss these expansions and to demonstrate by example the nature of the approximation of various truncations to some well-known density functions.

#### C.1 FORMAL DEVELOPMENT OF THE GRAM-CHARLIER EXPANSION

Consider a random variable  $\xi$  with a known density function and let  $x$  be the normalized random variable

$$x = \frac{\xi - m}{\sigma}$$

The  $m$  and  $\sigma$  are the mean and standard deviation associated with  $\xi$ . Denote the density function for  $x$  by  $f(x)$  and let  $\psi(x)$  represent the gaussian density function with mean zero and unit variance.

Consider an expansion of  $f(x)$  having the form

$$f(x) = c_0 \psi(x) + c_1 \psi'(x) + \frac{c_2}{2!} \psi''(x) + \dots \quad (C.1)$$

where  $\psi^{(k)}(x)$  is the  $k^{\text{th}}$  derivative of  $\psi$  and the  $c_k$  are the constants. The derivatives  $\psi^{(k)}$  are related to the Hermite polynomials according to [10]

$$\left(\frac{d}{dx}\right)^n e^{-x^2/2} = (-1)^n H_n(x) e^{-x^2/2} \quad (C.2)$$

The Hermite polynomials satisfy the orthogonality condition

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} H_m(x) H_n(x) e^{-x^2/2} dx = \begin{cases} n! & \text{for } m = n \\ 0 & \text{for } m \neq n \end{cases} \quad (C.3)$$

Thus, (C.1) is an expansion in orthogonal polynomials and the  $c_k$  can be determined from the orthogonality condition (C.3). Multiply  $f(x)$  by  $H_k(x)$  and integrate. It follows that

$$c_k = (-1)^k \int_{-\infty}^{\infty} H_k(x) f(x) dx \quad (C.4)$$

The Hermite polynomials can be established directly from (C.2). It is easily seen that

$$\begin{aligned} H_0(x) &= 1 \\ H_1(x) &= x \\ H_2(x) &= x^2 - 1 \\ H_3(x) &= x^3 - 3x \\ H_4(x) &= x^4 - 6x^2 + 3 \\ H_5(x) &= x^5 - 10x^3 + 15x \\ H_6(x) &= x^6 - 15x^4 + 45x^2 - 15 \\ &\vdots \\ &\vdots \\ &\vdots \end{aligned}$$

In general, the polynomials satisfy the recurrence relation

$$H_{n+1}(x) = xH_n(x) - nH_{n-1}(x) \quad (C.5)$$

From the Hermite polynomials and (C.4), the coefficients are found. In particular, one finds that

$$\begin{aligned} c_0 &= 1 \\ c_1 &= 0 \\ c_2 &= 0 \end{aligned}$$

$$c_3 = - \frac{\mu_3}{\sigma^3}$$

$$c_4 = \frac{\mu_4}{\sigma^4} - 3$$

$$c_5 = - \frac{\mu_5}{\sigma^5} + 10 \frac{\mu_3}{\sigma^3}$$

$$c_6 = \frac{\mu_6}{\sigma^6} - 15 \frac{\mu_4}{\sigma^4} + 30$$

.

.

.

The  $\mu_k$  are the central moments associated with the random variable  $\xi$ .

The coefficients  $c_k$  have been designated by the name "quasi-moments" by Stratonovich [57, 58]. It is interesting to observe that  $c_1$  and  $c_2$  are identically zero and that the first  $k$  quasi-moments are completely determined by the first  $k$  central moments. The central moments have a more commonly understood significance, and they are dealt with in the text rather than the  $c_k$ .

The Edgeworth expansion is closely related to the Gram-Charlier, but its derivation is somewhat more involved. It arose from considerations relating to random variables  $\xi$  which are given as the sum of  $n$  random variables  $\xi_1$

$$\xi = \xi_1 + \xi_2 + \dots + \xi_n$$

According to the central limit theorem (with the suitable restrictions), the  $\xi$  should be approximately gaussian when  $n$  is large. In this case, it is

desirable that all terms of the same order in  $n$  be included when the expansion is truncated. It was found that under this constraint, the expansion for the normalized variable  $x$  is

$$\begin{aligned}
 f(x) = & \psi(x) \\
 & + \frac{1}{3!} c_3 \psi^{(3)}(x) \\
 & + \frac{1}{4!} c_4 \psi^{(4)}(x) + \frac{10}{6!} c_3^2 \psi^{(6)}(x) \\
 & + \frac{1}{5!} c_5 \psi^{(5)}(x) + \frac{35}{6!} c_3 c_4 \psi^{(7)}(x) + \frac{280}{9!} c_3^3 \psi^{(9)}(x) \\
 & + \dots
 \end{aligned} \tag{C.6}$$

where terms of the same order in  $n$  are stated on the same line. The details regarding the derivation of (C.6) shall be omitted.

The Gram-Charlier and Edgeworth expansions can be written in terms of the characteristic functions of  $f(x)$  and  $\psi(x)$ . Let  $\varphi(s)$  be the characteristic function of  $f(x)$

$$\varphi(s) = \int_{-\infty}^{\infty} e^{isx} f(x) dx$$

The Fourier transform of the derivative of a gaussian density is given by

$$\int_{-\infty}^{\infty} e^{isx} \psi^{(k)}(x) dx = (-is)^k e^{-s^2/2}, \quad k = 0, 1, \dots \tag{C.7}$$

so the Gram-Charlier expansion becomes

$$\varphi(s) = e^{-s^2/2} \left[ 1 + \frac{c_3}{3!} (-is)^3 + \frac{c_4}{4!} (-is)^4 + \dots \right] \tag{C.8}$$

## C.2 APPROXIMATION OF DENSITIES USING THE EXPANSIONS

In this section we attempt to approximate densities that are distinctly nongaussian in order to examine the effects of various truncations of the expansions. First, consider a uniform distribution

$$p(\xi) = \begin{cases} \frac{1}{2h} & a - h \leq \xi \leq a + h \\ 0 & \text{elsewhere} \end{cases}$$

The moments of this distribution are known to be

$$\begin{aligned} m &= a \\ \sigma^2 &= h^2/3 \\ \mu_3 &= 0 = \mu_5 \\ \mu_4 &= h^4/5 \\ \mu_6 &= h^6/7 \end{aligned}$$

We shall only consider truncations of the expansions that contain at most the first six moments.

The density for the normalized variable is

$$f(x) = \begin{cases} \frac{1}{2\sqrt{3}} & -\sqrt{3} \leq x \leq \sqrt{3} \\ 0 & \text{elsewhere} \end{cases}$$

This density is symmetric about the mean, so the Gram-Charlier and Edgeworth expansions are identical through the sixth order moments. We shall consider the following approximations to  $f(x)$ .

$$\text{Case 1: } f(x) = \psi(x)$$

$$\text{Case 2: } f(x) = \psi(x) + \frac{1}{4!} c_4 \psi^{(4)}(x)$$

$$\text{Case 3: } f(x) = \psi(x) + \frac{1}{4!} c_4 \psi^{(4)}(x) + \frac{1}{6!} c_6 \psi^{(6)}(x)$$

The data pertaining to these approximations are contained in Figure (C-1). The important thing to observe about these plots is that Cases 2 and 3 exhibit negative values for the larger values of  $x$ . This is impossible for a true density, so it suggests a possible source of difficulty associated with the use of these approximations.

The second density to be considered is the  $\chi^2$  density.

$$p(\xi) = \begin{cases} \frac{1}{2^{n/2} \Gamma(n/2)} x^{n/2-1} e^{-x/2} & \text{for } x > 0 \\ 0 & \text{for } x \leq 0 \end{cases}$$

The moments  $\alpha_k$  (not the central moments) of this distribution are known to be given by

$$\alpha_k = E[\xi^k] = n(n+2) \dots (n+2k-2)$$

The central moments are easily determined.

$$m = n$$

$$\sigma^2 = 2n$$

$$\mu_3 = 8n$$

$$\mu_4 = 12n^2 + 48n$$

$$\mu_5 = 160n^2 + 384n$$

$$\mu_6 = 120n^3 + 2080n^2 + 3840n$$

The density for the associated normalized variable  $x$  is

$$f(x) = \begin{cases} \frac{(2n)^{1/2}}{2^{n/2} \Gamma(n/2)} (\sqrt{2n} x + n)^{n/2-1} e^{-\frac{\sqrt{2n} x + n}{2}} & x > -(n/2)^{1/2} \\ 0 & x \leq -(n/2)^{1/2} \end{cases}$$

This density is approximated for two values of  $n$ . Figure C-2 contains the results for  $n = 1$  and Figure C-3 depicts the approximations for  $n = 4$ .

These cases are plotted in each of these figures.

$$\text{Case 1: } f(x) = \psi(x) + \frac{1}{3!} c_3 \psi^{(3)}(x)$$

$$\text{Case 2: } f(x) = \psi(x) + \frac{1}{3!} c_3 \psi^{(3)}(x) + \frac{1}{4!} c_4 \psi^{(4)}(x)$$

$$\text{Case 3: } f(x) = \psi(x) + \frac{1}{3!} c_3 \psi^{(3)}(x) + \frac{1}{4!} c_4 \psi^{(4)}(x) + \frac{10}{6!} c_3^2 \psi^{(6)}(x)$$

The latter two have been included because they illustrate the differences between the approximations provided by the Gram-Charlier and by Edgeworth expansions. Both of these cases include the fourth order moments, but no moments of higher order. The Edgeworth expansion is seen to provide a significantly better approximation in each case.

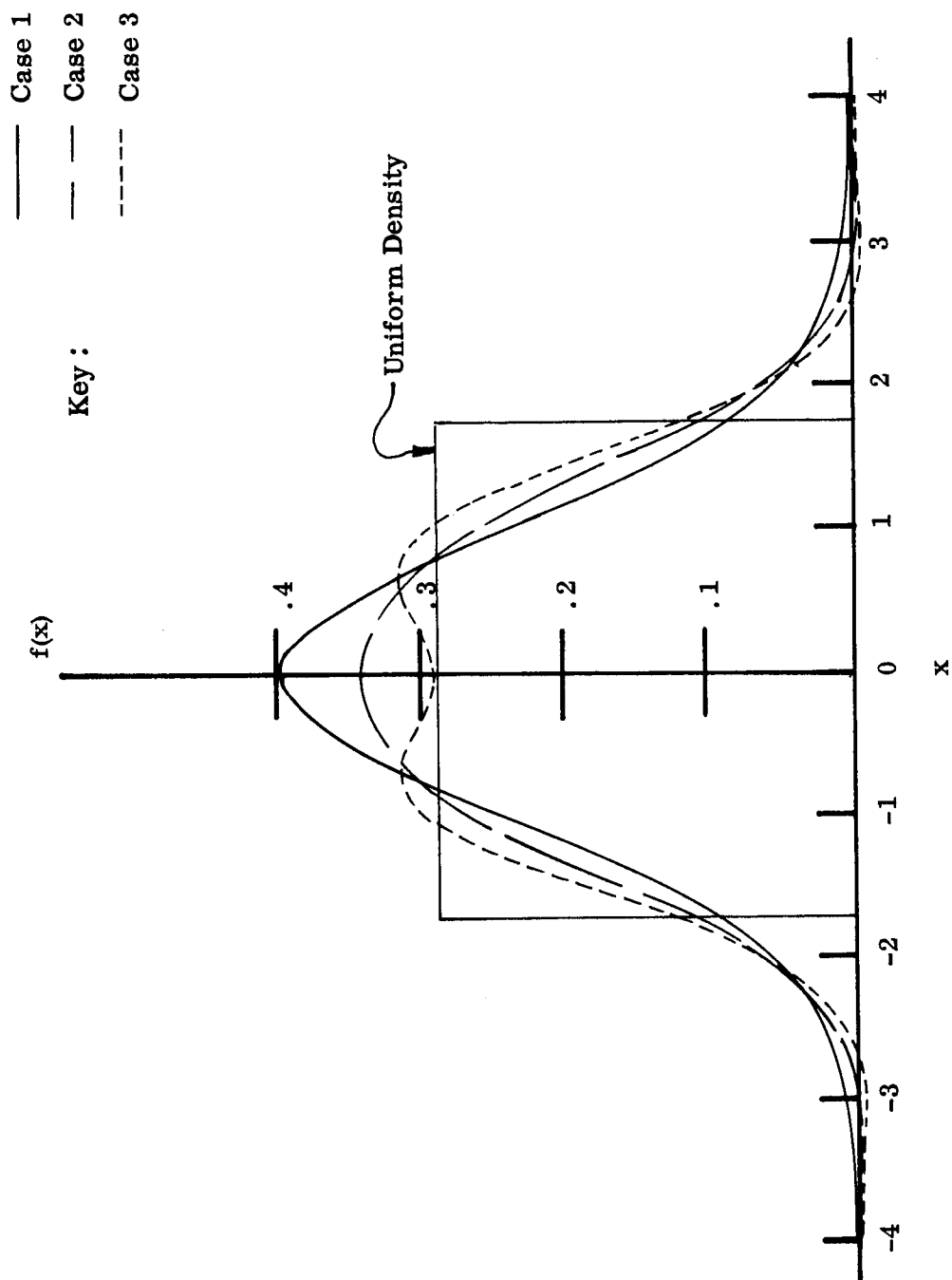


Figure C-1. Series Approximations to Uniform Probability Density



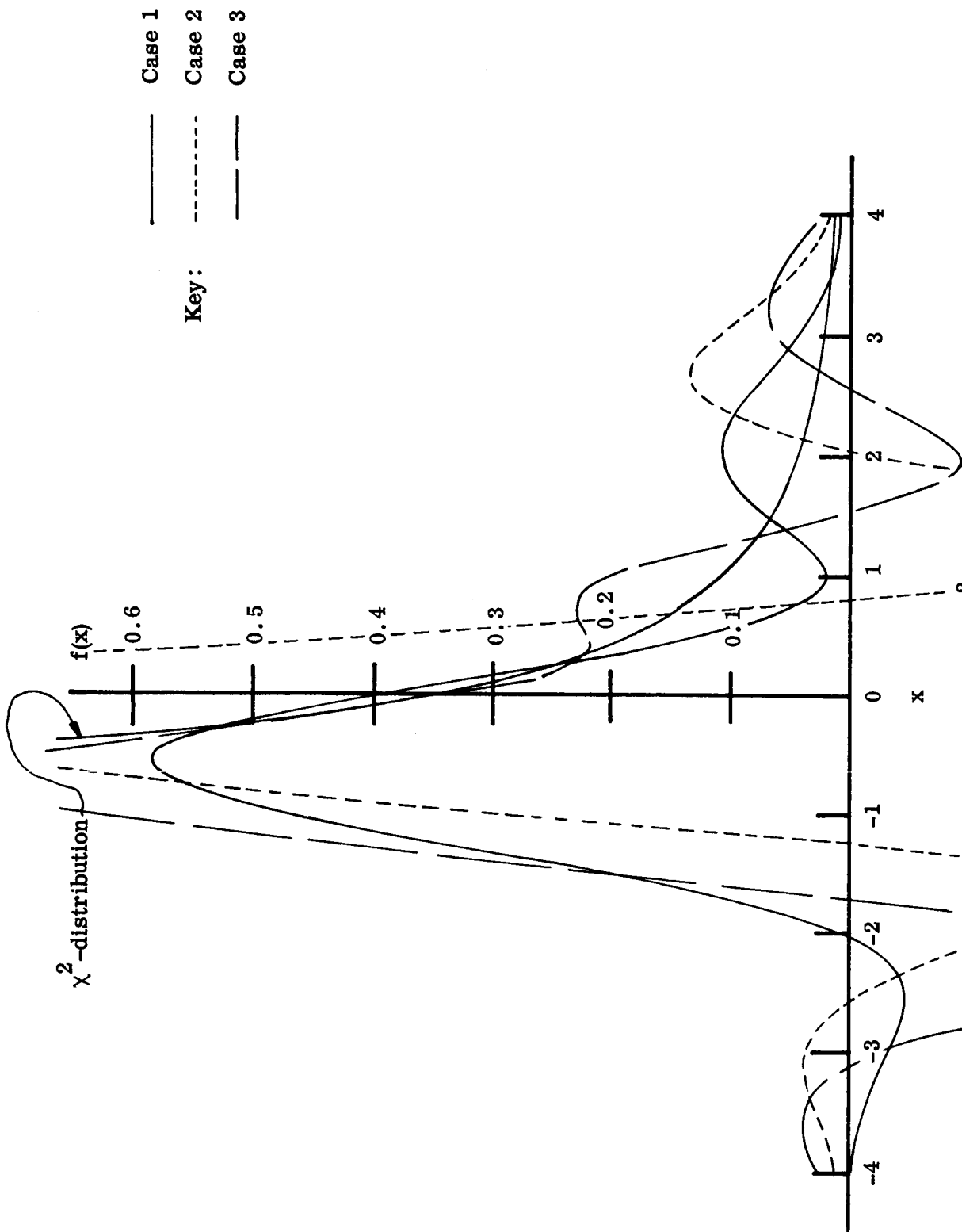


Figure C-2. Approximation of  $\chi^2$ -distribution for  $n = 1$

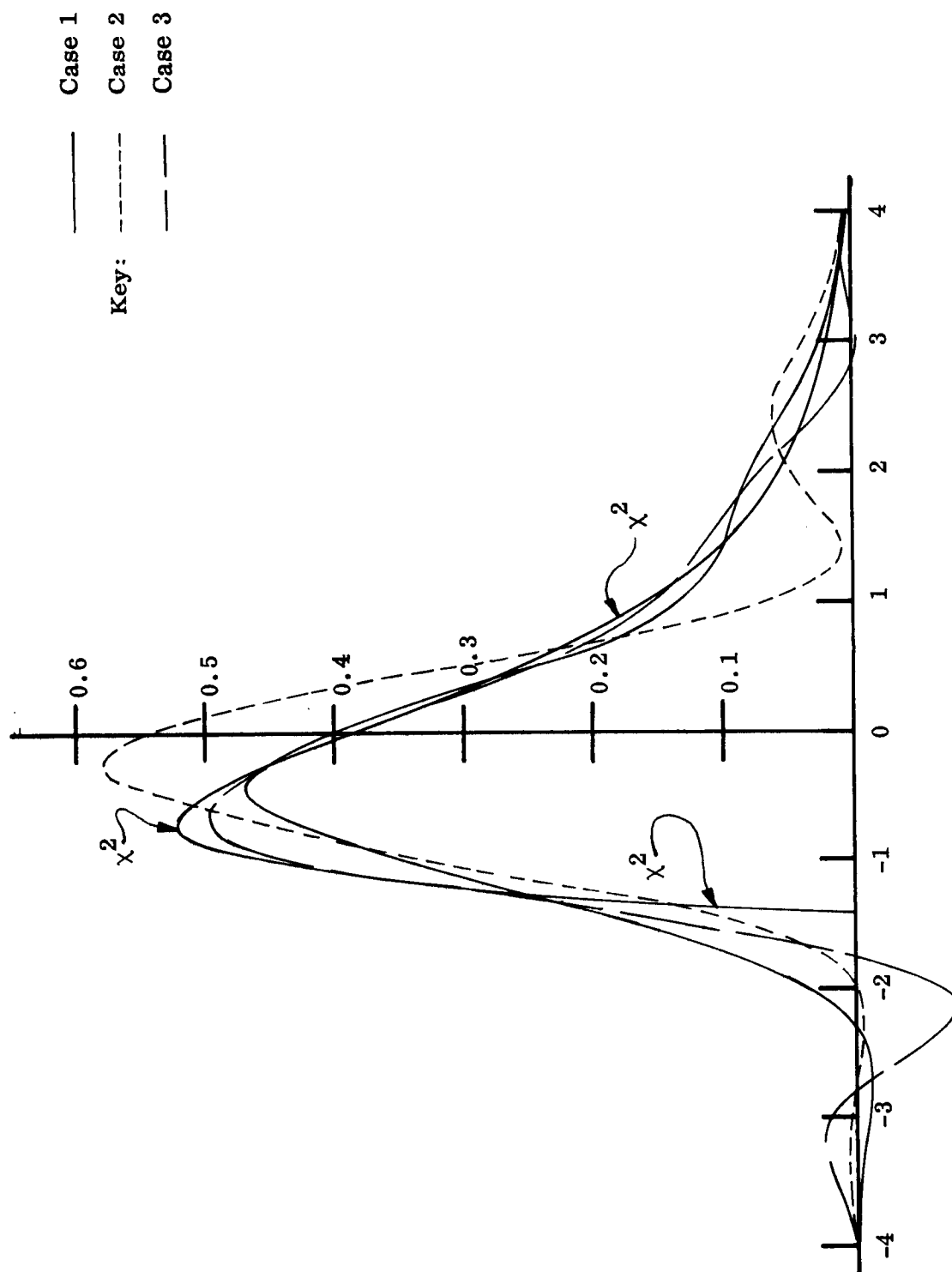


Figure C-3. Approximation of  $\chi^2$ -distribution for  $n = 4$

## APPENDIX D

### DERIVATION OF NONGAUSSIAN A POSTERIORI DENSITY FUNCTION

In this appendix equations defining the moments of the most general a posteriori density function considered in this investigation are derived. Attention is restricted to scalar plant and measurement equations. The techniques used in the derivation can be applied to more general density functions and/or more general system equations (including multidimensional systems). Increased generality leads only to algebraic difficulties, not conceptual difficulties.

Suppose that the plant and measurement equations are described by

$$x_k = f_k x_{k-1} + g_k x_{k-1}^2 + w_{k-1} \quad (D.1)$$

$$z_k = h_k x_k + e_k x_k^2 + v_k \quad (D.2)$$

where the additive noise  $\{w_k\}$  and  $\{v_k\}$  are white, gaussian sequences. The initial state  $x_0$  is also gaussian. The system (D.1) and (D.2) can be considered to represent the second-order Taylor series expansion of some more general nonlinear system. The a posteriori density shall be approximated at each sampling time by

$$p(x_k/z^k) = k_k \exp - \frac{1}{2} \zeta_k^2 \left[ 1 + \frac{1}{3!} c_3 H_3(\zeta_k) + \frac{1}{4!} c_4 H_4(\zeta_k) + \frac{10}{6!} c_3^2 H_6(\zeta_k) \right] \quad (D.3)$$

where

$$\zeta_k = \frac{x_k - \hat{x}_k}{p_k}$$

$$c_3 = - \frac{\mu_k}{3 p_k}$$

$$c_4 = \frac{\nu_k}{4 p_k} - 3$$

and

$$H_3(\zeta_k) = \zeta_k^3 - 3 \zeta_k$$

$$H_4(\zeta_k) = \zeta_k^4 - 6 \zeta_k^2 + 3$$

$$H_6(\zeta_k) = \zeta_k^6 - 15 \zeta_k^4 + 45 \zeta_k^2 - 1$$

The  $\hat{x}_k$  is the mean value and  $p_k^2$ ,  $\mu_k$ ,  $\nu_k$  are the second, third, and fourth central moments of  $p(x_k/z^k)$ . Relations defining the moments are derived below. In the derivation, use is made of general relations between the moments and central moments of a distribution and also of the special properties of the moments of a gaussian distribution. Relationships of this nature are summarized in Appendix E.

#### D.1 INITIAL SAMPLING TIME

Since the  $x_0$  is assumed to be gaussian and since a measurement  $z_0$  is assumed to be available, the initial sampling time is a special case of the more general results presented in Section D.3. According to (2.15)

$$p(x_0/z_0) = \frac{p(z_0/x_0)p(x_0)}{\int p(z_0/x_0)p(x_0)dx_0}$$

From (D.2), it is true that

$$\begin{aligned}
 p(z_o/x_o) &= k_v \exp - \frac{1}{2} \frac{(z_o - h_o x_o - e_o x_o^2)^2}{r_o^2} \\
 &= k_v \exp - \frac{1}{2} \frac{(z_o - h_o x_o)^2}{r_o^2} \\
 &\quad \exp - \frac{1}{2} \frac{(-2e_o z_o x_o^2 + 2h_o e_o x_o^3 + e_o^2 x_o^4)}{r_o^2}
 \end{aligned} \tag{D.4}$$

It is necessary to approximate the second exponential. Let

$$B(x_o) \stackrel{\text{Df}}{=} \frac{e_o z_o}{r_o^2} x_o^2 - \frac{h_o e_o}{r_o^2} x_o^3 - \frac{e_o^2}{2r_o^2} x_o^4$$

and let

$$b_2 \stackrel{\text{Df}}{=} \frac{e_o z_o}{r_o^2}$$

$$b_3 \stackrel{\text{Df}}{=} - \frac{h_o e_o}{r_o^2}$$

$$b_4 \stackrel{\text{Df}}{=} - \frac{e_o^2}{2r_o^2}$$

Before dealing further with the second factor, let us form  $p(x_o/z_o)$ . Let

$$k_o' \stackrel{\text{Df}}{=} \frac{1}{\int p(z_o/x_o) p(x_o) dx_o}$$

Then,

$$p(x_o/z_o) = k'_o k_v k_x \exp - \frac{1}{2} \left( \frac{x_o - a}{m_o} \right)^2 \exp - \frac{1}{2} \left( \frac{z_o - h_o x_o}{r_o^2} \right)^2 \exp B(x_o)$$

Combining the two gaussian terms and completing the square, one gets

$$p(x_o/z_o) = k'_o k_v k_x \exp - \frac{1}{2} \left( \frac{a^2}{m_o^2} + \frac{z_o^2}{r_o^2} - \frac{\hat{\xi}_o^2}{\pi_o^2} \right) \exp - \frac{1}{2} \left( \frac{x_o - \hat{\xi}_o}{\pi_o} \right)^2 \exp B(x_o)$$

where

$$\frac{1}{\pi_o^2} = \frac{h_o^2}{r_o^2} + \frac{1}{m_o^2}$$

$$\hat{\xi}_o = \pi_o^2 \left( \frac{h_o z_o}{r_o^2} + \frac{a}{m_o^2} \right)$$

The  $\hat{\xi}_o$  and  $\pi_o^2$  would be the conditional mean and variance if the measurements were linear. The factor  $\exp B(x_o)$  then serves as a correction to the linear results.

Now, rewrite  $B(x_o)$  in terms of the variable  $\hat{\xi}_o$ . That is, let

$$\begin{aligned} b_2 x_o^2 + b_3 x_o^3 + b_4 x_o^4 &= \beta_o + \beta_1 (x_o - \hat{\xi}_o) + \beta_{21} (x_o - \hat{\xi}_o)^2 \\ &+ \beta_{31} (x_o - \hat{\xi}_o)^3 + \beta_{41} (x_o - \hat{\xi}_o)^4 \end{aligned}$$

The values of the  $\beta_i$  that satisfy the equality are easily determined to be

$$\beta_o = b_2 \hat{\xi}_o^2 + b_3 \hat{\xi}_o^3 + b_4 \hat{\xi}_o^4$$

$$\beta_1 = 2b_2 \hat{\xi}_o + 3b_3 \hat{\xi}_o^2 + 4b_4 \hat{\xi}_o^3$$

$$\beta_{21} = b_2 + 3b_3 \hat{\xi}_o + 6b_4 \hat{\xi}_o^2$$

$$\beta_{31} = b_3 + 4b_4 \hat{\xi}_o$$

$$\beta_{41} = b_4$$

The constant  $\beta_o$  can be considered separately. Define  $B_1(\eta_o)$  as

$$B_1(\eta_o) \stackrel{\text{Df}}{=} \beta_1 \eta_o + \beta_{21} \eta_o^2 + \beta_{31} \eta_o^3 + \beta_{41} \eta_o^4 \quad (\text{D.5})$$

where

$$\eta_o \stackrel{\text{Df}}{=} x_o - \hat{\xi}_o$$

Now, let us approximate  $\exp B_1(\eta_o)$  by a power series. Then,

$$\exp B_1(\eta_o) = 1 + B_1(\eta_o) + \frac{1}{2!} [B_1(\eta_o)]^2 + o[B_1^3(\eta_o)]$$

Terms of greater than second order shall be neglected. Let

$$\begin{aligned} B_2(\eta_o) &\stackrel{\text{Df}}{=} \frac{1}{2} [B_1(\eta_o)]^2 \\ &= \beta_{22} \eta_o^2 + \beta_{32} \eta_o^3 + \beta_{42} \eta_o^4 + \beta_5 \eta_o^5 + \beta_6 \eta_o^6 + \beta_7 \eta_o^7 + \beta_8 \eta_o^8 \end{aligned}$$

where

$$\beta_{22} \stackrel{\text{Df}}{=} \beta_1^2 / 2$$

$$\beta_{32} \stackrel{\text{Df}}{=} \beta_1 \beta_{21}$$

$$\beta_{42} \stackrel{\text{Df}}{=} (\beta_{21}^2 + 2\beta_1 \beta_{31}) / 2$$

$$\beta_5 \stackrel{\text{Df}}{=} \beta_1 \beta_{41} + \beta_{21} \beta_{31}$$

$$\beta_6 \stackrel{\text{Df}}{=} (\beta_{31}^2 + 2\beta_{21}\beta_{41})/2$$

$$\beta_7 \stackrel{\text{Df}}{=} \beta_{31}\beta_{41}$$

$$\beta_8 \stackrel{\text{Df}}{=} \beta_{41}^2/2$$

Thus,

$$p(x_o/z_o) = k_o' k_v k_x \exp - \frac{1}{2} \left( \frac{a^2}{m_o^2} + \frac{z_o^2}{r_o^2} - \frac{\xi_o^2}{\pi_o^2} \right) \exp \beta_o \exp - \frac{1}{2} \left( \frac{\eta_o}{\pi_o} \right)^2 [1 + B_1(\eta_o) + B_2(\eta_o)]$$

The constant  $[k_v k_x \exp - 1/2 (a^2/m_o^2 + z_o^2/r_o^2 - \xi_o^2/\pi_o^2) \exp \beta_o]$  will be cancelled by the corresponding term in  $\int p(z_o/x_o)p(x_o)dx_o$ . Define the constant  $k_o$  to be

$$\begin{aligned} k_o &= k_o' k_v k_x \exp - \frac{1}{2} \left( \frac{a^2}{m_o^2} + \frac{z_o^2}{r_o^2} - \frac{\xi_o^2}{\pi_o^2} \right) \exp \beta_o \\ &= \frac{1}{\sqrt{2\pi} \pi_o} \left[ \frac{1}{\sqrt{2\pi} \pi_o} \int \exp - \frac{1}{2} \left( \frac{\eta_o}{\pi_o} \right)^2 [1 + B_1(\eta_o) + B_2(\eta_o)] d\eta_o \right]^{-1} \end{aligned}$$

Note the change of variable for the integration. Thus

$$k_o = \frac{1}{\sqrt{2\pi} \pi_o} \frac{1}{1 + E[B_1] + E[B_2]}$$

and

$$p(x_o/z_o) = \frac{1}{1 + E[B_1] + E[B_2]} \left[ \frac{1}{\sqrt{2\pi} \pi_o} \exp - \frac{1}{2} \left( \frac{\eta_o}{\pi_o} \right)^2 [1 + B_1(\eta_o) + B_2(\eta_o)] \right] \quad (\text{D.6})$$



The expected value is obviously taken relative to the gaussian density

$$1/\sqrt{2\pi} \pi_o \exp - 1/2(\eta_o/\pi_o)^2.$$

To cause  $p(x_o/z_o)$  to assume the form prescribed by (D.3), determine the moments  $E[\eta_o^i/z_o]$

$$\begin{aligned} E[(x_o - \hat{\xi}_o)^i/z_o] &= \int (x_o - \hat{\xi}_o)^i p(x_o/z_o) dx_o \\ &= \frac{1}{1 + E(B_1) + E(B_2)} \left[ \frac{1}{\sqrt{2\pi} \pi_o} \int [\eta_o^i + \eta_o^i B_1(\eta_o) \right. \\ &\quad \left. + \eta_o^i B_2(\eta_o)] \exp - \frac{1}{2} \left( \frac{\eta_o}{\pi_o} \right)^2 d\eta_o \right] \\ E[\eta_o^i/z_o] &= \frac{E[\eta_o^i] + E[\eta_o^i B_1(\eta_o)] + E[\eta_o^i B_2(\eta_o)]}{1 + E[B_1] + E[B_2]} \end{aligned} \quad (D.7)$$

The expectations indicated in (D.7) are easily determined.

$$\begin{aligned} E[B_1] &= E[\beta_{11}\eta_o + \beta_{21}\eta_o^2 + \beta_{31}\eta_o^3 + \beta_{41}\eta_o^4] \\ &= \beta_{21}\pi_o^2 + 3\beta_{41}\pi_o^4 \end{aligned}$$

$$\begin{aligned} E[B_2] &= E[\beta_{22}\eta_o^2 + \beta_{32}\eta_o^3 + \beta_{42}\eta_o^4 + \beta_5\eta_o^5 + \beta_6\eta_o^6 + \beta_7\eta_o^7 + \beta_8\eta_o^8] \\ &= \beta_{22}\pi_o^2 + 3\beta_{42}\pi_o^4 + 5\beta_6\pi_o^6 + 7\beta_8\pi_o^8 \end{aligned}$$

This follows from the fact [10] that for a gaussian variable  $\eta_o$  with mean zero and standard derivation  $\pi_o$

$$E[\eta_o^i] = \begin{cases} 0, & i = 1, 3, 5, \dots \\ (i-1) \pi_o^i, & i = 2, 4, 6, \dots \end{cases}$$

The  $E[\eta_0^i B_1(\eta_0)]$  and  $E[\eta_0^i B_2(\eta_0)]$  are determined in the same manner.

The central moments can be determined directly from the  $E[\eta_0^i/z_0]$ .

For example,

$$E[\eta_0/z_0] = E[(x_0 - \hat{\xi}_0)/z_0] = E[x_0/z_0] - \hat{\xi}_0$$

By definition,

$$\hat{x}_0 = E[x_0/z_0]$$

so

$$\hat{x}_0 = \hat{\xi}_0 + E[\eta_0/z_0] \quad (D.8)$$

Thus, the  $E(\eta_0/z_0)$  provides a correction to the conditional mean obtained from the linear density.

The variance is determined by considering the second moment.

$$\begin{aligned} E[\eta_0^2/z_0] &= E[(x_0 - \hat{\xi}_0)^2/z_0] \\ &= E\{[(x_0 - \hat{x}_0) + (\hat{x}_0 - \hat{\xi}_0)]^2/z_0\} \\ &= E[(x_0 - \hat{x}_0)^2/z_0] + 2E[(x_0 - \hat{x}_0)/z_0](\hat{x}_0 - \hat{\xi}_0) + (\hat{x}_0 - \hat{\xi}_0)^2 \end{aligned}$$

But

$$E[x_0 - \hat{x}_0/z_0] = \hat{x}_0 - \hat{x}_0 = 0.$$

and

$$\hat{x}_0 - \hat{\xi}_0 = E[\eta_0/z_0]$$

so

$$\begin{aligned} p_o^2 &\stackrel{\text{Df}}{=} E[(x_o - \hat{x}_o)^2 / z_o] \\ &= E[\eta_o^2 / z_o] - E^2[\eta_o / z_o] \end{aligned} \quad (\text{D.9})$$

The third and fourth central moments are found in the same manner and are given by

$$\begin{aligned} \mu_o &= E[(x_o - \hat{x}_o)^3 / z_o] \\ &= E[\eta_o^3 / z_o] - 3E[\eta_o / z_o] p_o^2 - E^3[\eta_o / z_o] \end{aligned} \quad (\text{D.10})$$

$$\begin{aligned} \nu_o &= E[(x_o - \hat{x}_o)^4 / z_o] \\ &= E[\eta_o^4 / z_o] - 4\mu_o E[\eta_o / z_o] - 6p_o^2 E^2[\eta_o / z_o] - E^4[\eta_o / z_o] \end{aligned} \quad (\text{D.11})$$

This completes the derivation of the moments for the first sampling time.

## D.2 THE PREDICTION DENSITY $p(x_k / z^{k-1})$

The density  $p(x_k / z^{k-1})$ , according to (2.13), is given in general by

$$p(x_k / z^{k-1}) = \int p(x_{k-1} / z^{k-1}) p(x_k / x_{k-1}) dx_{k-1}$$

It is necessary to proceed carefully in determining this density. In particular, if one attempts to establish  $p(x_k / z^{k-1})$  directly from the formula, it becomes apparent that one is led to a hopeless morass of algebraic manipulation. On the other hand, it has been pointed out that the object of the approximation procedure is to determine the moments of the distribution. That is, it is desired to determine

$$E[x_k^i / z^{k-1}] = \int x_k^i p(x_k / z^{k-1}) dx_k \quad (\text{D.12})$$

This can be written as

$$E[x_k^i/z^{k-1}] = \int x_k^i \{ \int p(x_{k-1}/z^{k-1}) p(x_k/x_{k-1}) dx_{k-1} \} dx_k$$

Iterating the integrals, this becomes

$$E[x_k^i/z^{k-1}] = \int p(x_{k-1}/z^{k-1}) \{ \int x_k^i p(x_k/x_{k-1}) dx_k \} dx_{k-1}$$

The innermost integration can be easily accomplished because of the assumptions on the plant. Consider the mean value.

$$E[x_k/x_{k-1}] = \int_{-\infty}^{\infty} x_k \exp - \frac{1}{2} \left[ \frac{x_k - (f_k x_{k-1} + g_k x_{k-1}^2)}{q_{k-1}} \right]^2 dx_k$$

It follows immediately that

$$E[x_k/x_{k-1}] = f_k x_{k-1} + g_k x_{k-1}^2$$

Thus,

$$E[x_k/z^{k-1}] = \int (f_k x_{k-1} + g_k x_{k-1}^2) p(x_{k-1}/z^{k-1}) dx_{k-1}$$

From the definition of  $p(x_{k-1}/z^{k-1})$ , one obtains immediately

$$E[x_k/z^{k-1}] = f_k \hat{x}_{k-1} + g_k (p_{k-1}^2 + \hat{x}_{k-1}^2)$$

$$\stackrel{\text{Df}}{=} \hat{x}_{k/k-1} \quad (\text{D. 13})$$

Continue in the same manner to determine the higher order moments. Thus,

$$E[x_k^2/x_{k-1}] = q_{k-1}^2 + (f_k x_{k-1} + g_k x_{k-1}^2)^2$$

and so

$$E[x_k^2/z^{k-1}] = q_{k-1}^2 + f_k^2 E[x_{k-1}^2/z^{k-1}] + 2f_k g_k E[x_{k-1}^3/z^{k-1}] + g_k^2 E[x_{k-1}^4/z^{k-1}]$$

By definition, one sees that

$$E[x_{k-1}^2/z^{k-1}] = p_{k-1}^2 + \hat{x}_{k-1}^2$$

$$E[x_{k-1}^3/z^{k-1}] = \mu_{k-1} + 3\hat{x}_{k-1}p_{k-1}^2 + \hat{x}_{k-1}^3$$

$$E[x_{k-1}^4/z^{k-1}] = \nu_{k-1} + 4\hat{x}_{k-1}\mu_{k-1} + 6\hat{x}_{k-1}^2p_{k-1}^2 + \hat{x}_{k-1}^4$$

The conditional variance is given by

$$p_{k/k-1}^2 = E[x_k^2/z^{k-1}] - \hat{x}_{k/k-1}^2$$

After some manipulation, this is found to be

$$\begin{aligned} p_{k/k-1}^2 &= q_{k-1}^2 + (f_k + 2g_k \hat{x}_{k-1})^2 p_{k-1}^2 \\ &\quad + 2g_k(f_k + 2g_k \hat{x}_{k-1})\mu_{k-1} + g_k^2(\nu_{k-1} - p_{k-1}^4) \end{aligned} \quad (D.14)$$

The third order moment is found in a similar manner.

$$E[x_k^3/x_{k-1}] = 3q_{k-1}^2(f_k x_{k-1} + g_k x_{k-1}^2) + (f_k x_{k-1} + g_k x_{k-1}^2)^3$$

so

$$\begin{aligned} E[x_k^3/z^{k-1}] &= 3q_{k-1}^2 \hat{x}_{k/k-1} + f_k^3 E[x_{k-1}^3/z^{k-1}] + 3f_k^2 g_k E[x_{k-1}^4/z^{k-1}] \\ &\quad + 3f_k g_k^2 E[x_{k-1}^5/z^{k-1}] + g_k^3 E[x_{k-1}^6/z^{k-1}] \end{aligned} \quad (D.15)$$

The first four moments  $E[x_{k-1}^i/z^{k-1}]$  ( $i = 1, 2, 3, 4$ ) are known from the determination of  $p(x_{k-1}/z^{k-1})$  and are used in the coefficients of the Edgeworth expansion. At the time that these moments are established, it would also be possible to determine as many of the higher order moments that are required to define  $p(x_k/z^{k-1})$ . Alternatively, one could view the truncation of the

Edgeworth expansion as implicitly requiring that the higher order moments be the same as the basic gaussian density. Then,

$$\mu_5 = \mu_7 = 0$$

$$\mu_6 = 5p_{k-1}^6$$

$$\mu_8 = 7p_{k-1}^8$$

This assumption will be used here. Then,

$$E[x_{k-1}^5/z^{k-1}] = 5\hat{x}_{k-1}v_{k-1} + 10\hat{x}_{k-1}^2\mu_{k-1} + 10\hat{x}_{k-1}^3p_{k-1}^2 + \hat{x}_{k-1}^5$$

$$E[x_{k-1}^6/z^{k-1}] = 5p_{k-1}^6 + 15\hat{x}_{k-1}^2v_{k-1} + 20\hat{x}_{k-1}^3\mu_{k-1}$$

$$+ 15\hat{x}_{k-1}^4p_{k-1} + \hat{x}_{k-1}^6$$

$$E[x_{k-1}^7/z^{k-1}] = 35\hat{x}_{k-1}p_{k-1}^6 + 35\hat{x}_{k-1}^3v_{k-1} + 35\hat{x}_{k-1}^4\mu_{k-1}$$

$$+ 21\hat{x}_{k-1}^5p_{k-1}^2 + \hat{x}_{k-1}^7$$

$$E[x_{k-1}^8/z^{k-1}] = 7p_{k-1}^8 + 140\hat{x}_{k-1}^2p_{k-1}^6 + 70\hat{x}_{k-1}^4v_{k-1} + 56\hat{x}_{k-1}^5\mu_{k-1}$$

$$+ 28\hat{x}_{k-1}^6p_{k-1}^2 + \hat{x}_{k-1}^8$$

The third central moment is related to the  $E[x_k^3/z^{k-1}]$  according to

$$\mu_{k/k-1} = E(x_k^3/z^{k-1}) - 3\hat{x}_{k/k-1}p_{k/k-1}^2 - \hat{x}_{k/k-1}^3$$

The analytic expression for the  $\mu_{k/k-1}$  shall be derived for future reference.

It is not necessary since (D.15), (D.16) and the relations for the moments

$E[x_{k-1}^i/z^{k-1}]$  can be used to determine  $\mu_{k/k-1}$ .

In particular, after considerable manipulation, one finds

$$\begin{aligned}
\mu_{k/k-1} = & f_k^3 \mu_{k-1} + 3f_k^2 g_k (2\hat{x}_{k-1} \mu_{k-1} + \nu_k - p_{k-1}^4) \\
& + 3f_k g_k^2 (11\hat{x}_{k-1} p_{k-1}^4 - \hat{x}_{k-1} \nu_{k-1} - 6\hat{x}_{k-1}^2 \mu_{k-1} - 2p_{k-1}^2 \mu_{k-1}) \\
& + g_k^3 (7p_{k-1}^6 + 33p_{k-1}^4 \hat{x}_{k-1}^2 - 3p_{k-1}^2 \nu_{k-1} - 12\hat{x}_{k-1} p_{k-1}^2 \mu_{k-1} \\
& - 3\nu_{k-1} \hat{x}_{k-1}^2 - 12\hat{x}_{k-1}^3 + \mu_{k-1}) \quad (D. 17)
\end{aligned}$$

The fourth order moment is determined from

$$E[x_k^4/x_{k-1}] = 3q_{k-1}^4 + 6(f_k x_{k-1} + g_k x_{k-1}^2)^2 q_{k-1}^2 + (f_k x_{k-1} + g_k x_{k-1}^2)^4$$

Thus,

$$\begin{aligned}
E(x_k^4/z^{k-1}) = & 3q_{k-1}^4 + 6q_{k-1}^2 [f_k^2 E(x_{k-1}^2/z^{k-1}) + 2f_k g_k E(x_{k-1}^3/z^{k-1}) + g_k^2 E(x_{k-1}^4/z^{k-1})] \\
& + f_k^4 E(x_{k-1}^4/z^{k-1}) + 4f_k^3 g_k E(x_{k-1}^5/z^{k-1}) + 6f_k^2 g_k^2 E(x_{k-1}^6/z^{k-1}) \\
& + 4f_k g_k^3 E(x_{k-1}^7/z^{k-1}) + g_k^4 E(x_{k-1}^8/z^{k-1}) \quad (D. 18)
\end{aligned}$$

The fourth central moment is determined from  $E(x_k^4/z^{k-1})$

$$\nu_{k/k-1} = E(x_k^4/z^{k-1}) - 4\hat{x}_{k/k-1} \mu_{k/k-1} - 6\hat{x}_{k/k-1}^2 p_{k/k-1}^2 - \hat{x}_{k/k-1} \quad (D. 19)$$

Equations (D. 18) and (D. 19) serve to define  $\nu_{k/k-1}$ .

This completes the derivation of  $p(x_k/z^{k-1})$ . Recall that the only approximation is in the number of terms retained in the expansion and in the moments of greater than fourth order.

### D.3 THE GENERAL RELATION FOR $p(x_k/z^k)$

The general relations for the moments of  $p(x_k/z^k)$  are derived in a manner similar to that utilized in Section D.1. The density  $p(x_k/z^{k-1})$  has the form given by (D.3) rather than the gaussian  $p(x_0)$ , so the derivation is somewhat more complicated algebraically.

According to (IV), the  $p(x_k/z^k)$  is described by

$$p(x_k/z^k) = \frac{p(x_k/z^{k-1})p(z_k/x_k)}{p(z_k/z^{k-1})}$$

Let

$$k'_k \stackrel{\text{Df}}{=} \frac{1}{p(z_k/z^{k-1})}$$

Then

$$p(x_k/z^k) = k'_k k_{k/k-1} k_v \exp - \frac{1}{2} \left( \frac{z_k - k_k x_k - e_k x_k^2}{r_k} \right)^2 \exp - \frac{1}{2} (\zeta_{k/k-1})^2$$

$$\left[ 1 + \frac{1}{3!} c_3 H_3(\zeta_{k/k-1}) + \frac{1}{4!} c_4 H_4(\zeta_{k/k-1}) + \frac{10}{6!} c_3^2 H_6(\zeta_{k/k-1}) \right] \quad (\text{D.20})$$

where

$$\zeta_{k/k-1} = \left( \frac{x_k - \hat{x}_{k/k-1}}{p_{k/k-1}} \right)$$

This can be rewritten as



$$\begin{aligned}
p(x_k/z^k) &= k! k_{k/k-1} k_v \exp - \frac{1}{2} \left( \frac{\hat{x}_{k/k-1}^2}{p_{k/k-1}} + \frac{z_k^2}{r_k} - \frac{\hat{\xi}_k^2}{\pi_k} \right) \exp \beta_0 \\
&\exp - \frac{1}{2} \left( \frac{x_k - \hat{\xi}_k}{\pi_k} \right)^2 [1 + B_1(\eta_k) + B_2(\eta_k)] \\
&[1 + \frac{1}{3!} c_3 H_3(\eta_k) + \frac{1}{4!} c_4 H_4(\eta_k) + \frac{10}{6!} c_3^2 H_6(\eta_k)] \quad (D. 21)
\end{aligned}$$

where

$$\begin{aligned}
\frac{1}{\pi_k} &= \frac{h_k^2}{r_k} + \frac{1}{p_{k/k-1}} \\
\hat{\xi}_k &= \pi_k^2 \left( \frac{h_k z_k}{r_k^2} + \frac{\hat{x}_{k/k-1}}{p_{k/k-1}} \right)
\end{aligned}$$

The  $B_1(\eta_k)$  and  $B_2(\eta_k)$  are defined in the preceding section with the trivial change of subscript. The Hermite polynomials can be rewritten in terms of the variable

$$\eta_k = (x_k - \hat{\xi}_k).$$

First, the Hermite polynomials can be written as

$$\begin{aligned}
H_3(\zeta_{k/k-1}) &= s_0 + s_1 x_k + s_2 x_k^2 + s_3 x_k^3 \\
H_4(\zeta_{k/k-1}) &= d_0 + d_1 x_k + d_2 x_k^2 + d_3 x_k^3 + d_4 x_k^4 \\
H_6(\zeta_{k/k-1}) &= e_0 + e_1 x_k + e_2 x_k^2 + e_3 x_k^3 + e_4 x_k^4 + e_5 x_k^5 + e_6 x_k^6
\end{aligned}$$

where

$$s_0 \stackrel{\text{Df}}{=} - \left( \frac{\hat{x}_{k/k-1}^3}{3 p_{k/k-1}} - \frac{3 \hat{x}_{k/k-1}}{p_{k/k-1}} \right)$$

$$s_1 \stackrel{\text{Df}}{=} 3 \left( \frac{\hat{x}_{k/k-1}^2}{3 p_{k/k-1}} - \frac{1}{p_{k/k-1}} \right)$$

$$s_2 \stackrel{\text{Df}}{=} - 3 \frac{\hat{x}_{k/k-1}}{3 p_{k/k-1}}$$

$$s_3 \stackrel{\text{Df}}{=} \frac{1}{3 p_{k/k-1}}$$

and

$$d_0 \stackrel{\text{Df}}{=} \frac{\hat{x}_{k/k-1}^4}{4 p_{k/k-1}} - 6 \frac{\hat{x}_{k/k-1}^2}{2 p_{k/k-1}} + 3$$

$$d_1 \stackrel{\text{Df}}{=} - 4 \left( \frac{\hat{x}_{k/k-1}^3}{4 p_{k/k-1}} - 3 \frac{\hat{x}_{k/k-1}}{2 p_{k/k-1}} \right)$$

$$d_2 \stackrel{\text{Df}}{=} 6 \left( \frac{\hat{x}_{k/k-1}^2}{4 p_{k/k-1}} - \frac{1}{2 p_{k/k-1}} \right)$$

$$d_3 \stackrel{\text{Df}}{=} - 4 \frac{\hat{x}_{k/k-1}}{4 p_{k/k-1}}$$

$$d_4 \stackrel{\text{Df}}{=} \frac{1}{4 p_{k/k-1}}$$

and finally,

$$e_0 \stackrel{\text{Df}}{=} \frac{\hat{x}_{k/k-1}^6}{p_{k/k-1}} - 15 \frac{\hat{x}_{k/k-1}^4}{p_{k/k-1}} + 45 \frac{\hat{x}_{k/k-1}^2}{p_{k/k-1}} - 15$$

$$e_1 \stackrel{\text{Df}}{=} -6 \left( \frac{\hat{x}_{k/k-1}^5}{p_{k/k-1}} - 10 \frac{\hat{x}_{k/k-1}^3}{p_{k/k-1}} + 15 \frac{\hat{x}_{k/k-1}}{p_{k/k-1}} \right)$$

$$e_2 \stackrel{\text{Df}}{=} 15 \left( \frac{\hat{x}_{k/k-1}^4}{p_{k/k-1}} - 6 \frac{\hat{x}_{k/k-1}^2}{p_{k/k-1}} + 3 \frac{1}{p_{k/k-1}} \right)$$

$$e_3 \stackrel{\text{Df}}{=} -20 \left( \frac{\hat{x}_{k/k-1}^3}{p_{k/k-1}} - 3 \frac{\hat{x}_{k/k-1}}{p_{k/k-1}} \right)$$

$$e_4 \stackrel{\text{Df}}{=} 15 \left( \frac{\hat{x}_{k/k-1}^2}{p_{k/k-1}} - \frac{1}{p_{k/k-1}} \right)$$

$$e_5 \stackrel{\text{Df}}{=} -6 \frac{\hat{x}_{k/k-1}}{p_{k/k-1}}$$

$$e_6 = \frac{1}{p_{k/k-1}}$$

These polynomials are to be rewritten as

$$H_3(\eta_k) = \psi_0 + \psi_1 \eta_k + \psi_2 \eta_k^2 + \psi_3 \eta_k^3$$

$$H_4(\eta_k) = \delta_0 + \delta_1 \eta_k + \delta_2 \eta_k^2 + \delta_3 \eta_k^3 + \delta_4 \eta_k^4$$

$$H_6(\eta_k) = \epsilon_0 + \epsilon_1 \eta_k + \epsilon_2 \eta_k^2 + \epsilon_3 \eta_k^3 + \epsilon_4 \eta_k^4 + \epsilon_5 \eta_k^5 + \epsilon_6 \eta_k^6$$

The coefficients that satisfy the equality are found to be

$$\psi_0 = s_0 + s_1 \hat{\xi}_k + s_2 \hat{\xi}_k^2 + s_3 \hat{\xi}_k^3$$

$$\psi_1 = s_1 + 2s_2 \hat{\xi}_k + 3s_3 \hat{\xi}_k^2$$

$$\psi_2 = s_2 + 3s_3 \hat{\xi}_k$$

$$\psi_3 = s_3$$

$$\delta_0 = d_0 + d_1 \hat{\xi}_k + d_2 \hat{\xi}_k^2 + d_3 \hat{\xi}_k^3 + d_4 \hat{\xi}_k^4$$

$$\delta_1 = d_1 + 2d_2 \hat{\xi}_k + 3d_3 \hat{\xi}_k^2 + 4d_4 \hat{\xi}_k^3$$

$$\delta_2 = d_2 + 3d_3 \hat{\xi}_k + 6d_4 \hat{\xi}_k^2$$

$$\delta_3 = d_3 + 4d_4 \hat{\xi}_k$$

$$\delta_4 = d_4$$

$$\epsilon_0 = e_0 + e_1 \hat{\xi}_k + e_2 \hat{\xi}_k^2 + e_3 \hat{\xi}_k^3 + e_4 \hat{\xi}_k^4 + e_5 \hat{\xi}_k^5 + e_6 \hat{\xi}_k^6$$

$$\epsilon_1 = e_1 + 2e_2 \hat{\xi}_k + 3e_3 \hat{\xi}_k^2 + 4e_4 \hat{\xi}_k^3 + 5e_5 \hat{\xi}_k^4 + 6e_6 \hat{\xi}_k^5$$

$$\epsilon_2 = e_2 + 3e_3 \hat{\xi}_k + 6e_4 \hat{\xi}_k^2 + 10e_5 \hat{\xi}_k^3 + 15e_6 \hat{\xi}_k^4$$

$$\epsilon_3 = e_3 + 4e_4 \hat{\xi}_k + 10e_5 \hat{\xi}_k^2 + 20e_6 \hat{\xi}_k^3$$

$$\epsilon_4 = e_4 + 5e_5 \hat{\xi}_k + 15e_6 \hat{\xi}_k^2$$

$$\epsilon_5 = e_5 + 6e_6 \hat{\xi}_k$$

$$\epsilon_6 = e_6$$

Let

$$\begin{aligned}
 k_o &\stackrel{\text{Df}}{=} k'_k v_{k/k-1} \exp - \frac{1}{2} \left( \frac{\hat{x}_{k/k-1}}{p_{k/k-1}} + \frac{z_k^2}{r_k} - \frac{\xi_k^2}{\pi_k} \right) \exp \beta_o \\
 &= \frac{1}{\sqrt{2\pi} \pi_o} \left[ \frac{1}{\sqrt{2\pi} \pi_o} \int \exp - \frac{1}{2} \left( \frac{\eta_k}{\pi_k} \right)^2 [1 + B_1 + B_2] \right. \\
 &\quad \left. [1 + \frac{1}{3!} c_3 H_3 + \frac{1}{4!} c_4 H_4 + \frac{10}{6!} c_3^2 H_6] d\eta_k \right]^{-1}
 \end{aligned}$$

Then,

$$\begin{aligned}
 p(x_k/z^k) &= \frac{1}{E\{[1+B_1+B_2][1+\frac{1}{3!}c_3H_3+\frac{1}{4!}c_4H_4+\frac{10}{6!}c_3^2H_6]\}} \\
 &\quad \left\{ - \frac{1}{\sqrt{2\pi} \pi_o} \int \exp - \frac{1}{2} \left( \frac{\eta_k}{\pi_k} \right)^2 [1 + B_1 + B_2] \right. \\
 &\quad \left. [1 + \frac{1}{3!} c_3 H_3 + \frac{1}{4!} c_4 H_4 + \frac{10}{6!} c_3^2 H_6] \right\} \quad (D.22)
 \end{aligned}$$

The moments  $E[\eta_k^i/z^k]$  are computed in a straightforward manner. The products of the polynomial terms appearing in (D.22) are required. These products are stated below for reference.

$$\begin{aligned}
 B_1 H_3 &= s_o \beta_1 \eta_k + (s_o \beta_{21} + s_1 \beta_1) \eta_k^2 + (s_o \beta_{31} + s_1 \beta_{21} + s_2 \beta_1) \eta_k^3 \\
 &\quad + (s_o \beta_{41} + s_1 \beta_{31} + s_2 \beta_{21} + s_3 \beta_1) \eta_k^4 + (s_1 \beta_{41} + s_2 \beta_{31} + s_3 \beta_{21}) \eta_k^5 \\
 &\quad + (s_2 \beta_{41} + s_3 \beta_{31}) \eta_k^6 + s_3 \beta_{41} \eta_k^7
 \end{aligned}$$

$$\begin{aligned}
 B_1 H_4 &= d_o \beta_1 \eta_k + (d_o \beta_{21} + d_1 \beta_1) \eta_k^2 + (d_o \beta_{31} + d_1 \beta_{21} + d_2 \beta_1) \eta_k^3 \\
 &\quad + (d_o \beta_{41} + d_1 \beta_{31} + d_2 \beta_{21} + d_3 \beta_1) \eta_k^4 + (d_1 \beta_{41} + d_2 \beta_{31} + d_3 \beta_{21} + d_4 \beta_1) \eta_k^5 \\
 &\quad + (d_2 \beta_{41} + d_3 \beta_{31} + d_4 \beta_{21}) \eta_k^6 + (d_3 \beta_{41} + d_4 \beta_{31}) \eta_k^7 + d_4 \beta_{41} \eta_k^8
 \end{aligned}$$

$$\begin{aligned}
B_1 H_6 = & \beta_1 e_o \eta_k + (\beta_1 e_1 + \beta_{21} e_o) \eta_k^2 + (\beta_1 e_2 + \beta_{21} e_1 + \beta_{31} e_o) \eta_k^3 \\
& + (\beta_1 e_3 + \beta_{21} e_2 + \beta_{31} e_1 + \beta_{41} e_o) \eta_k^4 + (\beta_1 e_4 + \beta_{21} e_3 + \beta_{31} e_2 + \beta_{41} e_1) \eta_k^5 \\
& + (\beta_1 e_5 + \beta_{21} e_4 + \beta_{31} e_3 + \beta_{41} e_2) \eta_k^6 + (\beta_1 e_6 + \beta_{21} e_5 + \beta_{31} e_4 + \beta_{41} e_3) \eta_k^7 \\
& + (\beta_{21} e_6 + \beta_{31} e_5 + \beta_{41} e_4) \eta_k^8 + (\beta_{31} e_6 + \beta_{41} e_5) \eta_k^9 + \beta_{41} e_6 \eta_k^{10}
\end{aligned}$$

$$\begin{aligned}
B_2 H_3 = & s_o \beta_{22} \eta_k^2 + (s_1 \beta_{22} + s_o \beta_{32}) \eta_k^3 + (s_o \beta_{42} + s_1 \beta_{32} + s_2 \beta_{22}) \eta_k^4 \\
& + (s_o \beta_5 + s_1 \beta_{42} + s_2 \beta_{32} + s_3 \beta_{22}) \eta_k^5 + (s_o \beta_6 + s_1 \beta_5 + s_2 \beta_{42} + s_3 \beta_{32}) \eta_k^6 \\
& + (s_o \beta_7 + s_1 \beta_6 + s_2 \beta_5 + s_3 \beta_{42}) \eta_k^7 + (s_o \beta_8 + s_1 \beta_7 + s_2 \beta_6 + s_3 \beta_5) \eta_k^8 \\
& + (s_1 \beta_8 + s_2 \beta_7 + s_3 \beta_6) \eta_k^9 + (s_2 \beta_8 + s_3 \beta_7) \eta_k^{10} + s_3 \beta_8 \eta_k^{11}
\end{aligned}$$

$$\begin{aligned}
B_2 H_4 = & d_o \beta_{22} \eta_k^2 + (d_o \beta_{32} + d_1 \beta_{22}) \eta_k^3 + (d_o \beta_{42} + d_1 \beta_{32} + d_2 \beta_{22}) \eta_k^4 \\
& + (d_o \beta_5 + d_1 \beta_{42} + d_2 \beta_{32} + d_3 \beta_{22}) \eta_k^5 + (d_o \beta_6 + d_1 \beta_5 + d_2 \beta_{42} + d_3 \beta_{32} + d_4 \beta_{22}) \eta_k^6 \\
& + (d_o \beta_7 + d_1 \beta_6 + d_2 \beta_5 + d_3 \beta_{42} + d_4 \beta_{32}) \eta_k^7 + (d_o \beta_8 + d_1 \beta_7 + d_2 \beta_6 + d_3 \beta_5 + d_4 \beta_{42}) \eta_k^8 \\
& + (d_1 \beta_8 + d_2 \beta_7 + d_3 \beta_6 + d_4 \beta_5) \eta_k^9 + (d_2 \beta_8 + d_3 \beta_7 + d_4 \beta_6) \eta_k^{10} \\
& + (d_3 \beta_8 + d_4 \beta_7) \eta_k^{11} + (d_4 \beta_8) \eta_k^{12}
\end{aligned}$$

$$\begin{aligned}
B_2 H_6 = & e_o \beta_{22} \eta_k^2 + (e_1 \beta_{22} + e_o \beta_{32}) \eta_k^3 + (e_o \beta_{42} + e_1 \beta_{32} + e_2 \beta_{22}) \eta_k^4 \\
& + (e_o \beta_5 + e_1 \beta_{42} + e_2 \beta_{32} + e_3 \beta_{22}) \eta_k^5 + (e_o \beta_6 + e_1 \beta_5 + e_2 \beta_{42} + e_3 \beta_{32} + e_4 \beta_{22}) \eta_k^6 \\
& + (e_o \beta_7 + e_1 \beta_6 + e_2 \beta_5 + e_3 \beta_{42} + e_4 \beta_{32} + e_5 \beta_{22}) \eta_k^7 \\
& + (e_o \beta_8 + e_1 \beta_7 + e_2 \beta_6 + e_3 \beta_5 + e_4 \beta_{42} + e_5 \beta_{32} + e_6 \beta_{22}) \eta_k^8 \\
& + (e_1 \beta_8 + e_2 \beta_7 + e_3 \beta_6 + e_4 \beta_5 + e_5 \beta_{42} + e_6 \beta_{32}) \eta_k^9 \\
& + (e_2 \beta_8 + e_3 \beta_7 + e_4 \beta_6 + e_5 \beta_5 + e_6 \beta_{42}) \eta_k^{10} + (e_3 \beta_8 + e_4 \beta_7 + e_5 \beta_6 + e_6 \beta_5) \eta_k^{11} \\
& + (e_4 \beta_8 + e_5 \beta_7 + e_6 \beta_6) \eta_k^{12} + (e_5 \beta_8 + e_6 \beta_7) \eta_k^{13} + e_6 \beta_8 \eta_k^{14}
\end{aligned}$$

After the moments  $E[\eta_k^i/z^k]$  ( $i = 1, 2, 3, 4$ ) have been computed, the central moments required for the Edgeworth expansion are computed from (D.8) through (D.11) with the obvious change of subscript. This completes the derivation.

## APPENDIX E

### MOMENTS OF A DISTRIBUTION

The relation between the moments and the central moments are presented in this appendix to provide a convenient reference. The central moments through the twentieth are included.

#### 5.1 MOMENTS OF A GENERAL DISTRIBUTION

Consider a random variable  $\xi$  with finite moments of all order. The central moments are related to the moments according to the following schedule.

$$\begin{aligned}
 E[\xi] &= a \\
 E[(\xi - a)^2] &\stackrel{\text{Df}}{=} \sigma^2 \\
 &= E[\xi^2] - a^2 \\
 E[(\xi - a)^3] &\stackrel{\text{Df}}{=} \mu_3 \\
 &= E[\xi^3] - 3a\sigma^2 - a^3 \\
 E[(\xi - a)^4] &= \mu_4 \\
 &= E[\xi^4] - 4a\mu_3 - 6a^2\sigma^2 - a^4 \\
 E[(\xi - a)^5] &\stackrel{\text{Df}}{=} \mu_5 \\
 &= E[\xi^5] - 5a\mu_4 - 10a^2\mu_3 - 10a^3\sigma^2 - a^5 \\
 E[(\xi - a)^6] &\stackrel{\text{Df}}{=} \mu_6 \\
 &= E[\xi^6] - 6a\mu_5 - 15a^2\mu_4 - 20a^3\mu_3 - 15a^4\sigma^2 - a^6 \\
 E[(\xi - a)^7] &\stackrel{\text{Df}}{=} \mu_7 \\
 &= E[\xi^7] - 7a\mu_6 - 21a^2\mu_5 - 35a^3\mu_4 - 35a^4\mu_3 - 21a^5\sigma^2 - a^7
 \end{aligned}$$



$$E[(\xi - a)^8] \stackrel{\text{Df}}{=} \mu_8$$

$$= E[\xi^8] - 8a\mu_7 - 28a^2\mu_6 - 56a^3\mu_5 - 70a^4\mu_4 - 56a^5\mu_3 - 28a^6\sigma^2 - a^8$$

$$E[(\xi - a)^9] \stackrel{\text{Df}}{=} \mu_9$$

$$= E[\xi^9] - 9a\mu_8 - 36a^2\mu_7 - 84a^3\mu_6 - 126a^4\mu_5 - 126a^5\mu_4 - 84a^6\mu_3 - 36a^7\sigma^2 - a^9$$

$$E[(\xi - a)^{10}] \stackrel{\text{Df}}{=} \mu_{10}$$

$$= E[\xi^{10}] - 10a\mu_9 - 45a^2\mu_8 - 120a^3\mu_7 - 210a^4\mu_6 - 252a^5\mu_5 - 210a^6\mu_4 - 120a^7\mu_3 - 45a^8\sigma^2 - a^{10}$$

$$E[(\xi - a)^{11}] \stackrel{\text{Df}}{=} \mu_{11}$$

$$E[\xi^{11}] - 11a\mu_{10} - 55a^2\mu_9 - 165a^3\mu_8 - 330a^4\mu_7 - 462a^5\mu_6 - 462a^6\mu_5 - 330a^7\mu_4 - 165a^8\mu_3 - 55a^9\sigma^2 - a^{11}$$

$$E[(\xi - a)^{12}] \stackrel{\text{Df}}{=} \mu_{12}$$

$$= E[\xi^{12}] - 12\mu_{11}a - 66\mu_{10}a^2 - 220\mu_9a^3 - 495\mu_8a^4 - 792\mu_7a^5 - 924\mu_6a^6 - 792\mu_5a^7 - 495\mu_4a^8 - 220\mu_3a^9 - 66\mu_2a^{10} - a^{12}$$

$$E[(\xi - a)^{13}] \stackrel{\text{Df}}{=} \mu_{13}$$

$$= E[\xi^{13}] - 13\mu_{12}a - 78\mu_{11}a^2 - 286\mu_{10}a^3 - 715\mu_9a^4 - 1287\mu_8a^5 - 1716\mu_7a^6 - 1716\mu_6a^7 - 1287\mu_5a^8 - 715\mu_4a^9 - 286\mu_3a^{10} - 78\mu_2a^{11} - a^{13}$$

$$E[(\xi - a)^{14}] \stackrel{\text{Df}}{=} \mu_{14}$$

$$= E[\xi^{14}] - 14\mu_{13}a - 91\mu_{12}a^2 - 364\mu_{11}a^3 - 1001\mu_{10}a^4 - 2002\mu_9a^5 - 3003\mu_8a^6 - 3432\mu_7a^7 - 3003\mu_6a^8 - 2002\mu_5a^9 - 1001\mu_4a^{10} - 364\mu_3a^{11} - 91\mu_2a^{12} - a^{14}$$

$$\begin{aligned}
E[(\xi - a)^{15}] &\stackrel{\text{Df}}{=} \mu_{15} \\
&= E[\xi^{15}] - 15\mu_{14}a - 105\mu_{13}a^2 - 455\mu_{12}a^3 - 1365\mu_{11}a^4 \\
&\quad - 3003\mu_{10}a^5 - 5005\mu_9a^6 - 6435\mu_8a^7 - 6435\mu_7a^8 \\
&\quad - 5005\mu_6a^9 - 3003\mu_5a^{10} - 1365\mu_4a^{11} - 455\mu_3a^{12} \\
&\quad - 105\mu_2a^{13} - a^{15}
\end{aligned}$$

$$\begin{aligned}
E[(\xi - a)^{16}] &\stackrel{\text{Df}}{=} \mu_{16} \\
&= E[\xi^{16}] - 16\mu_{15}a - 120\mu_{14}a^2 - 560\mu_{13}a^3 - 1820\mu_{12}a^4 \\
&\quad - 4368\mu_{11}a^5 - 8008\mu_{10}a^6 - 11440\mu_9a^7 - 12870\mu_8a^8 \\
&\quad - 11440\mu_7a^9 - 8008\mu_6a^{10} - 4368\mu_5a^{11} - 1820\mu_4a^{12} \\
&\quad - 560\mu_3a^{13} - 120\mu_2a^{14} - a^{16}
\end{aligned}$$

$$\begin{aligned}
E[(\xi - a)^{17}] &\stackrel{\text{Df}}{=} \mu_{17} \\
&= E[\xi^{17}] - 17\mu_{16}a - 136\mu_{15}a^2 - 680\mu_{14}a^3 - 2380\mu_{13}a^4 \\
&\quad - 6188\mu_{12}a^5 - 12376\mu_{11}a^6 - 19448\mu_{10}a^7 - 24310\mu_9a^8 \\
&\quad - 24310\mu_8a^9 - 19448\mu_7a^{10} - 12376\mu_6a^{11} - 6188\mu_5a^{12} \\
&\quad - 2380\mu_4a^{13} - 680\mu_3a^{14} - 136\mu_2a^{15} - a^{17}
\end{aligned}$$

$$\begin{aligned}
E[(\xi - a)^{18}] &\stackrel{\text{Df}}{=} \mu_{18} \\
&= E[\xi^{18}] - 18\mu_{17}a - 153\mu_{16}a^2 - 816\mu_{15}a^3 - 3060\mu_{14}a^4 \\
&\quad - 8568\mu_{13}a^5 - 18564\mu_{12}a^6 - 31824\mu_{11}a^7 \\
&\quad - 43758\mu_{10}a^8 - 48620\mu_9a^9 - 43758\mu_8a^{10} \\
&\quad - 31824\mu_7a^{11} - 18564\mu_6a^{12} - 8568\mu_5a^{13} \\
&\quad - 3060\mu_4a^{14} - 816\mu_3a^{15} - 153\mu_2a^{16} - a^{18}
\end{aligned}$$

$$\begin{aligned}
E[(\xi - a)^{19}] &\stackrel{\text{Df}}{=} \mu_{19} \\
&= E[\xi^{19}] - 19\mu_{19}a - 171\mu_{17}a^2 - 969\mu_{16}a^3 - 3876\mu_{15}a^4 \\
&\quad - 11628\mu_{14}a^5 - 27132\mu_{13}a^6 - 50388\mu_{12}a^7 \\
&\quad - 75582\mu_{11}a^8 - 92378\mu_{10}a^9 - 92378\mu_9a^{10} \\
&\quad - 75582\mu_8a^{11} - 50388\mu_7a^{12} - 27132\mu_6a^{13} \\
&\quad - 11628\mu_5a^{14} - 3876\mu_4a^{15} - 969\mu_3a^{16} \\
&\quad - 171\sigma^2a^{17} - a^{19}
\end{aligned}$$

$$\begin{aligned}
E[(\xi - a)^{20}] &\stackrel{\text{Df}}{=} \mu_{20} \\
&= E[\xi^{20}] - 20\mu_{19}a - 190\mu_{18}a^2 - 1140\mu_{17}a^3 - 4845\mu_{16}a^4 \\
&\quad - 15504\mu_{15}a^5 - 38760\mu_{14}a^6 - 77520\mu_{13}a^7 \\
&\quad - 125970\mu_{12}a^8 - 167960\mu_{11}a^9 - 184756\mu_{10}a^{10} \\
&\quad - 167960\mu_9a^{11} - 125970\mu_8a^{12} - 77520\mu_7a^{13} \\
&\quad - 38760\mu_6a^{14} - 15504\mu_5a^{15} - 4845\mu_4a^{16} \\
&\quad - 1140\mu_3a^{17} - 190\sigma^2a^{18} - a^{20}
\end{aligned}$$

## E.2 MOMENTS OF A GAUSSIAN DISTRIBUTION

The central moments of a gaussian distribution have the properties that

$$\mu_i = 0, \quad i = 3, 5, 7, 9, \dots$$

$$\mu_i = (i-1)\sigma^i, \quad i = 4, 6, 8, \dots$$

Thus, the relations of the preceding section can be simplified for this distribution. These relations shall be stated explicitly below because they are of considerable importance in the approximations presented elsewhere in this document.

Let  $\xi$  be a gaussian random variable with mean value  $a$  and variance  $\sigma^2$ .

$$\begin{aligned}
 E[\xi] &= a \\
 E[\xi^2] &= \sigma^2 + a^2 \\
 E[\xi^3] &= a(3\sigma^2 + a^2) \\
 E[\xi^4] &= 3\sigma^4 + 6a^2\sigma^2 + a^4 \\
 E[\xi^5] &= a(15\sigma^4 + 10\sigma^2a^2 + a^4) \\
 E[\xi^6] &= 5\sigma^6 + 45\sigma^4a^2 + 15\sigma^2a^4 + a^6 \\
 E[\xi^7] &= a(35\sigma^6 + 105\sigma^4a^2 + 21\sigma^2a^4 + a^6) \\
 E[\xi^8] &= 7\sigma^8 + 140\sigma^6a^2 + 210\sigma^4a^4 + 28\sigma^2a^6 + a^8 \\
 E[\xi^9] &= a(63\sigma^8 + 420\sigma^6a^2 + 378\sigma^4a^4 + 36\sigma^2a^6 + a^8) \\
 E[\xi^{10}] &= 9\sigma^{10} + 315\sigma^8a^2 + 1050\sigma^6a^4 + 630\sigma^4a^6 + 45\sigma^2a^8 + a^{10} \\
 E[\xi^{11}] &= a(99\sigma^{10} + 1155\sigma^8a^2 + 2310\sigma^6a^4 + 990\sigma^4a^6 + 55\sigma^2a^8 + a^{10}) \\
 E[\xi^{12}] &= 11\sigma^{12} + 594\sigma^{10}a^2 + 3465\sigma^8a^4 + 4620\sigma^6a^6 + 1485\sigma^4a^8 \\
 &\quad + 66\sigma^2a^{10} + a^{12} \\
 E[\xi^{13}] &= a(143\sigma^{12} + 2574\sigma^{10}a^2 + 9009\sigma^8a^4 + 8580\sigma^6a^6 + 2145\sigma^4a^8 \\
 &\quad + 78\sigma^2a^{10} + a^{12}) \\
 E[\xi^{14}] &= 13\sigma^{14} + 1001\sigma^{12}a^2 + 9009\sigma^{10}a^4 + 21021\sigma^8a^6 + 15015\sigma^6a^8 \\
 &\quad + 3003\sigma^4a^{10} + 91\sigma^2a^{12} + a^{14} \\
 E[\xi^{15}] &= a(195\sigma^{14} + 5005\sigma^{12}a^2 + 27027\sigma^{10}a^4 + 45045\sigma^8a^6 \\
 &\quad + 25025\sigma^6a^8 + 4095\sigma^4a^{10} + 105\sigma^2a^{12} + a^{14}) \\
 E[\xi^{16}] &= 15\sigma^{16} + 1560\sigma^{14}a^2 + 20020\sigma^{12}a^4 + 72072\sigma^{10}a^6 \\
 &\quad + 90090\sigma^8a^8 + 40040\sigma^6a^{10} + 5460\sigma^4a^{12} \\
 &\quad + 120\sigma^2a^{14} + a^{16}
 \end{aligned}$$

$$E[\xi^{17}] = a(255\sigma^{16} + 8840\sigma^{14}a^2 + 68068\sigma^{12}a^4 + 175032\sigma^{10}a^6 \\ + 170170\sigma^8a^8 + 61880\sigma^6a^{10} + 7140\sigma^4a^{12} \\ + 136\sigma^2a^{14} + a^{16})$$

$$E[\xi^{18}] = 17\sigma^{18} + 2295\sigma^{16}a^2 + 39780\sigma^{14}a^4 + 204204\sigma^{12}a^6 \\ + 393822\sigma^{10}a^8 + 306306\sigma^8a^{10} + 92820\sigma^6a^{12} \\ + 9180\sigma^4a^{14} + 153\sigma^2a^{16} + a^{18}$$

$$E[\xi^{19}] = a(323\sigma^{18} + 14535\sigma^{16}a^2 + 151164\sigma^{14}a^4 + 554268\sigma^{12}a^6 \\ + 831402\sigma^{10}a^8 + 529074\sigma^8a^{10} + 135660\sigma^6a^{12} \\ + 11628\sigma^4a^{14} + 171\sigma^2a^{16} + a^{18})$$

$$E[\xi^{20}] = 19\sigma^{20} + 3230\sigma^{18}a^2 + 72675\sigma^{16}a^4 + 503880\sigma^{14}a^6 \\ + 1385670\sigma^{12}a^8 + 1662804\sigma^{10}a^{10} + 881790\sigma^8a^{12} \\ + 193800\sigma^6a^{14} + 14535\sigma^4a^{16} + 190\sigma^2a^{18} + a^{20}$$